

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 30, 2005, 09:36:25 ; Search time 2698 Seconds  
(without alignments)  
18724.377 Million cell updates/sec

Title: US-10-650-249-1  
Perfect score: 7580  
Sequence: 1 agctctattaattcaagaga.....gccgcacccgggtggagct 7580

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues  
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*  
1: geneseqn1980s.\*  
2: geneseqn1990s.\*  
3: geneseqn2000s.\*  
4: geneseqn2001as.\*  
5: geneseqn2001bs.\*  
6: geneseqn2002as.\*  
7: geneseqn2002bs.\*  
8: geneseqn2003as.\*  
9: geneseqn2003bs.\*  
10: geneseqn2003cs.\*  
11: geneseqn2003ds.\*  
12: geneseqn2004as.\*  
13: geneseqn2004bs.\*  
14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7580	100.0	7580	12	ADL71829 Arabidops
2	1730	22.8	1746	12	ADL71846 Arabidops
3	1402.4	18.5	10078	6	ABQ73047 Tomato an
4	1367.4	18.0	17511	14	ADV39061 Plant gen
5	1367.4	18.0	18987	14	ADV39062 Plant gen
6	1360	17.9	1361	3	AAa88401 4x Camv 3
7	1235	16.3	1235	12	ADL71845 Arabidops
8	1062	14.0	1062	12	ADL71848 Arabidops
9	1058	14.0	1058	12	ADL71847 Arabidops
10	888	11.7	888	12	ADL71840 Arabidops
11	692	9.1	795	10	ADD30377 Plant yie
12	692	9.1	795	12	ADL44304 Plant tra
13	684	9.0	684	12	ADL71843 Arabidops
14	621	8.2	5033	2	AAQ74695 pPHI1406
15	617.2	8.1	4727	14	AEA40942 T-DNA reg
16	617.2	8.1	8428	8	ABT16592 Artificia
17	617.2	8.1	8428	10	ACC44694 Plasmid p
18	617.2	8.1	10122	8	ABT16598 Artificia
19	617.2	8.1	10122	10	ACC44701 Plasmid p

20	617.2	8.1	10549	8	ABT16593
21	617.2	8.1	10549	10	ACC44696
22	617.2	8.1	11182	8	ABT16591
23	617.2	8.1	11182	10	ACC44693
24	617.2	8.1	11478	8	ABV75873
25	617.2	8.1	11667	10	ADH89347
26	617.2	8.1	12304	8	ABV75876
27	617.2	8.1	12368	13	ADV16236
28	617.2	8.1	12497	8	ABV75875
29	617.2	8.1	12592	8	ABT16596
30	617.2	8.1	12592	10	ACC44699
31	617.2	8.1	12614	4	AAC66931
32	617.2	8.1	12921	3	AAa97523
33	617.2	8.1	14627	8	ABT16610
34	617.2	8.1	14627	10	ACC44712
35	616.8	8.1	829	4	AAF81265
36	616.8	8.1	829	9	ADA44822
37	615	8.1	5475	8	ABV75872
38	614.6	8.1	1334	2	AAZ27627
39	612.8	8.1	5013	14	AEA40941
40	612	8.1	4727	14	AEA40942
41	611.8	8.1	1216	8	ABV75874
42	610.6	8.1	781	14	AEA28179
43	610.6	8.1	781	14	AEA33917
44	599.6	7.9	11182	8	ABT16591
45	599.6	7.9	11182	10	ACC44693

ALIGNMENTS

RESULT 1  
ADL71829  
ID ADL71829 standard; DNA; 7580 BP.  
XX  
AC ADL71829;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Arabidopsis thaliana OBP3 DNA.  
XX  
KW Transgenic plant; Dof transcription factor; ocs binding factor;  
KW plant size; plant stature; root growth; plant; gene; ds; OBF;  
KW OBF binding protein; OBP3; SOBI; mouse-ear cross.  
XX  
OS Arabidopsis thaliana.  
XX  
FH Key Location/Qualifiers  
FT CDS 4103..5119  
FT /\*tag= a  
FT /product= "Arabidopsis thaliana OBP3 protein"  
FT /transl\_except= (pos:4811..4909, aa:Gly-Gly)  
XX  
PN US2004045055-A1.  
PD 04-MAR-2004.  
PP 28-AUG-2003; 2003US-00650249.  
XX  
PR 28-AUG-2002; 2002US-0406657P.  
XX  
(UNIW ) UNIV WASHINGTON.  
PI Neff MM;  
XX  
XX WPI; 2004-225757/21.  
XX P-PSDB; ADL71830.  
PT New transgenic plant cell, useful in producing plants with altered size  
XX and stature and with normal and healthy root growth.  
XX Claim 24; SEQ ID NO 1; 53pp; English.  
PS

Abt16593 Artificia  
Acc44696 Plasmid p  
Abt16591 Artificia  
Acc44693 Vector pa  
Abv75873 Vector pp  
Adh89347 Suppressi  
Abv75876 Luciferas  
Adv16236 Planc tra  
Abv75875 Luciferas  
Abt16596 Artificia  
Acc44699 Vector pa  
Aac66931 Plant sig  
Aaa97523 Plasmid p  
Abt16610 Artificia  
Acc44712 Plasmid p  
Aaf81265 Cauliflow  
Ada44822 Cauliflow  
Abv75872 Vector pt  
Aaz27627 Plasmid S  
Aea40941 T-DNA reg  
Aea40942 T-DNA reg  
Abv75874 Plant vec  
Aea28179 Alfalfa m  
Aea33917 Camv 35S  
Abt16591 Artificia  
Acc44693 Vector pa





Qy	4201	TGCAGGTTCCAGTCAAGCTAGAGTGAATTCAAATGGTGGAAAGTCTCGATCGCAAAAGT	4260
Db	4201	TGCAGGTTCCAGTCAAGCTAGAGTGAATTCAAATGGTGGAAAGTCTCGATCGCAAAAGT	4260
Qy	4261	CCCAATGCCCTGAAGCAGCTCTAAATGGCCCTAGATGTGACTCAACCAATCTAAAGTTCTG	4320
Db	4261	CCCAATGCCCTGAAGCAGCTCTAAATGGCCCTAGATGTGACTCAACCAATCTAAAGTTCTG	4320
Qy	4321	TTACTTTCAATAACTATAGCCTTACTCAACCTCGCCATTTCTCGAAAAATCATCTCGCGCTA	4380
Db	4321	TTACTTTCAATAACTATAGCCTTACTCAACCTCGCCATTTCTCGAAAAATCATCTCGCGCTA	4380
Qy	4381	TTGGACAGCTGGCGGTTCTTTCAGGAATGTTCTCTGTGGAGGAGGCTTTAGGAGGAACAA	4440
Db	4381	TTGGACAGCTGGCGGTTCTTTCAGGAATGTTCTCTGTGGAGGAGGCTTTAGGAGGAACAA	4440
Qy	4441	GAGAAGCAATCCAGATCGAAATCTACCGTCTGTGTCTCGACTGATATACTACTAGTAC	4500
Db	4441	GAGAAGCAATCCAGATCGAAATCTACCGTCTGTGTCTCGACTGATATACTACTAGTAC	4500
Qy	4501	TTCAATCACTTACTTTCGCGCCAAAGTTACTCAAAACCTTAGCAAGTTTCATAGCTACGGTCA	4560
Db	4501	TTCAATCACTTACTTTCGCGCCAAAGTTACTCAAAACCTTAGCAAGTTTCATAGCTACGGTCA	4560
Qy	4561	AATCCCGGAGTTAAATTCCAACTTGGCCACTTGTCTCTCTCCAAAGCCTTGGAGATTA	4620
Db	4561	AATCCCGGAGTTAAATTCCAACTTGGCCACTTGTCTCTCTCCAAAGCCTTGGAGATTA	4620
Qy	4621	CAATTCGAAGCAACACTGGATTTAGATTTTGGTGGAACTCAATAAGCAACATGATAAGTGG	4680
Db	4621	CAATTCGAAGCAACACTGGATTTAGATTTTGGTGGAACTCAATAAGCAACATGATAAGTGG	4680
Qy	4681	TATGAGTTCTAGTGGTGGATCTTTGATGCAATGGGAATACCTCCATCACAAAGCTCA	4740
Db	4681	TATGAGTTCTAGTGGTGGATCTTTGATGCAATGGGAATACCTCCATCACAAAGCTCA	4740
Qy	4741	GCAATTCCTCTTCTGCATCAACACTACCGGATTTGGTGGCAATCTTCAACGGGTTATATCC	4800
Db	4741	GCAATTCCTCTTCTGCATCAACACTACCGGATTTGGTGGCAATCTTCAACGGGTTATATCC	4800
Qy	4801	ATTACTAGAAAGTAAAGGAGGTGTTAATCAAGTGATTTCTCAACAGAAGAGTAGTGATTA	4860
Db	4801	ATTACTAGAAAGTAAAGGAGGTGTTAATCAAGTGATTTCTCAACAGAAGAGTAGTGATTA	4860
Qy	4861	TTCCAATCAGCTAATGTTTAAAGCCCTTGATGGATTTTCTTCAAGCGGGGTTAGCGGCAC	4920
Db	4861	TTCCAATCAGCTAATGTTTAAAGCCCTTGATGGATTTTCTTCAAGCGGGGTTAGCGGCAC	4920
Qy	4921	GCAACGAAGAATGTGAAGCGGAAGAGATGATCAGGATCGGGGTAGGGATGG	4980
Db	4921	GCAACGAAGAATGTGAAGCGGAAGAGATGATCAGGATCGGGGTAGGGATGG	4980
Qy	4981	AGTGAATAACTTATCAAGAACTTTTGGGTAAATCAACATAAACTCAGGCAGGAACGA	5040
Db	4981	AGTGAATAACTTATCAAGAACTTTTGGGTAAATCAACATAAACTCAGGCAGGAACGA	5040
Qy	5041	GGAAATACACATCATGGGGAGGTAAAGTTCTTGGAACCGGTTTCACTCCAACAACCTCAAC	5100
Db	5041	GGAAATACACATCATGGGGAGGTAAAGTTCTTGGAACCGGTTTCACTCCAACAACCTCAAC	5100
Qy	5101	AGGCCATCTCTCATTTAAAGTACTCAGCACTAGCTATTCTTCGATGATTTCTTTGTTGGTT	5160
Db	5101	AGGCCATCTCTCATTTAAAGTACTCAGCACTAGCTATTCTTCGATGATTTCTTTGTTGGTT	5160
Qy	5161	GGGGTGTACATTTGGTGTCTGTCTGCGAGTTATTGCTGAGGAAGATCAAAACCATCGAGCT	5220
Db	5161	GGGGTGTACATTTGGTGTCTGTCTGCGAGTTATTGCTGAGGAAGATCAAAACCATCGAGCT	5220
Qy	5221	ATATCCAAAGGCTAAATTTGAGGCTCAAGGAAGGTATGGTTATATAAACTATCTTTTGG	5280
Db	5221	ATATCCAAAGGCTAAATTTGAGGCTCAAGGAAGGTATGGTTATATAAACTATCTTTTGG	5280
Qy	5281	ATCTTTTAAAAAGATCTTCAAAAGTGTGAGTATGTTTATTTGGTTCGGTCTCTGGTATATTTA	5340
Db	5281	ATCTTTTAAAAAGATCTTCAAAAGTGTGAGTATGTTTATTTGGTTCGGTCTCTGGTATATTTA	5340
Qy	5341	TGTTTTTATTAGAAATTTGGTCTTATATATTTGGCTATATATAGAGGTGTGGGTGATATGTAT	5400
Db	5341	TGTTTTTATTAGAAATTTGGTCTTATATATTTGGCTATATATAGAGGTGTGGGTGATATGTAT	5400
Qy	5401	GAAATTCAGAGGTGATGTGGAAACTTTTGTGTGTTCATTTGAATAATCATCGAAATCTCT	5460
Db	5401	GAAATTCAGAGGTGATGTGGAAACTTTTGTGTGTTCATTTGAATAATCATCGAAATCTCT	5460
Qy	5461	CAATTTCTGGAGAGCCCAATATAGACATTTGAGACATCTATAGAACATATATGTAATGTA	5520
Db	5461	CAATTTCTGGAGAGCCCAATATAGACATTTGAGACATCTATAGAACATATATGTAATGTA	5520
Qy	5521	TATTTAAAAGCTTAATTAAGTCGAATTTTATGACCAAGTAAATAAATATGCGGAATGTACA	5580
Db	5521	TATTTAAAAGCTTAATTAAGTCGAATTTTATGACCAAGTAAATAAATATGCGGAATGTACA	5580
Qy	5581	TGCTAATATCGAGTTTAAACTATTTTTCCAAATAAACAATAATTTCTCTTTTCGTCCAA	5640
Db	5581	TGCTAATATCGAGTTTAAACTATTTTTCCAAATAAACAATAATTTCTCTTTTCGTCCAA	5640
Qy	5641	CTTATATACCTTATTTCTGATTTCTTATTTCTCTTTTAAATTTCTTTTCTTTTCTTTTCCAA	5700
Db	5641	CTTATATACCTTATTTCTGATTTCTTATTTCTCTTTTAAATTTCTTTTCTTTTCTTTTCCAA	5700
Qy	5701	GACACAAAAAAGAAAAATACAGAAACGAAAAAAGAGATTTTAAAAAATTCATAAACCCAC	5760
Db	5701	GACACAAAAAAGAAAAATACAGAAACGAAAAAAGAGATTTTAAAAAATTCATAAACCCAC	5760
Qy	5761	GAGAAATATGCACTTAAATTCAGACTAATCCGCCAAAATTTTCAGAAAAATTTATGTAATTTTG	5820
Db	5761	GAGAAATATGCACTTAAATTCAGACTAATCCGCCAAAATTTTCAGAAAAATTTATGTAATTTTG	5820
Qy	5821	CGAATTAATATTTGTTTCACAATCATATATGCGCCAACTAACTAATTTGAAAAAGACAATGAA	5880
Db	5821	CGAATTAATATTTGTTTCACAATCATATATGCGCCAACTAACTAATTTGAAAAAGACAATGAA	5880
Qy	5881	TGACTGAAACCATGATATCTCTCAAGTCTCAACCTATGAAGAATCATGTAAACCAATAG	5940
Db	5881	TGACTGAAACCATGATATCTCTCAAGTCTCAACCTATGAAGAATCATGTAAACCAATAG	5940
Qy	5941	ACTATCATCATGATTTAGTTAATGATGATCTATAATGTATTTCTTTGAACATAGATGTC	6000
Db	5941	ACTATCATCATGATTTAGTTAATGATGATCTATAATGTATTTCTTTGAACATAGATGTC	6000
Qy	6001	ATTTATCTGGATATAAAGATGGCGTTTAAACCTACTTTGCAATTTTGTGTTATATCTTTCT	6060
Db	6001	ATTTATCTGGATATAAAGATGGCGTTTAAACCTACTTTGCAATTTTGTGTTATATCTTTCT	6060
Qy	6061	TCTAATACATATGATCAATACACTTTTGTGTTTAAAAAGAAATTTAAAACTTATTTCAAC	6120
Db	6061	TCTAATACATATGATCAATACACTTTTGTGTTTAAAAAGAAATTTAAAACTT	



Db 6361 CTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAAATCCGAGAGGTTTCCCG 6420  
QY 6421 ATATTACCCTTTTGTGAAAGTCTCAATAGCCCTCTGGTCTTCTTGAGACTGTATCTTTGA 6480  
Db 6421 ATATTACCCTTTTGTGAAAGTCTCAATAGCCCTCTGGTCTTCTTGAGACTGTATCTTTGA 6480  
QY 6481 TATTCTTGAGTAGACGAGAGTGTGCTGTCCACCAATGTTGGGATCTAGATATCAATC 6540  
Db 6481 TATTCTTGAGTAGACGAGAGTGTGCTGTCCACCAATGTTGGGATCTAGATATCAATC 6540  
QY 6541 AATCCACTTGTCTTGAAGACGTGGTTGGAACGTCTTCTTTTCCAGATGTTCCCTGG 6600  
Db 6541 AATCCACTTGTCTTGAAGACGTGGTTGGAACGTCTTCTTTTCCAGATGTTCCCTGG 6600  
QY 6601 GTGGGGTCCATCTTTTGGGACCACTGTGCGTAGAGGCACTTTGAACGATAGCCTTTCC 6660  
Db 6601 GTGGGGTCCATCTTTTGGGACCACTGTGCGTAGAGGCACTTTGAACGATAGCCTTTCC 6660  
QY 6661 TATCGCAATGATGGCAATTTGTAGAACCAATCTCTTCTTCTACTGTCTTTCGATGAAGT 6720  
Db 6661 TATCGCAATGATGGCAATTTGTAGAACCAATCTCTTCTTCTACTGTCTTTCGATGAAGT 6720  
QY 6721 GACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATACCCCTTGTGAAAG 6780  
Db 6721 GACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATACCCCTTGTGAAAG 6780  
QY 6781 TCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTTGATATCTTGGAGTAGAGAGAG 6840  
Db 6781 TCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTTGATATCTTGGAGTAGAGAGAG 6840  
QY 6841 TGTGCTGCTCCACCAATGTTGGGATCTAGATATCAATCAATCCACTTGTGAAAGAG 6900  
Db 6841 TGTGCTGCTCCACCAATGTTGGGATCTAGATATCAATCAATCCACTTGTGAAAGAG 6900  
QY 6901 TGGTTGGACGCTCTTCTTTTCCAGATGTTCTGCTGGTGGGGTCCATCTTTGGGAC 6960  
Db 6901 TGGTTGGACGCTCTTCTTTTCCAGATGTTCTGCTGGTGGGGTCCATCTTTGGGAC 6960  
QY 6961 CACTGTGCGTAGAGGCACTCTTGAACGATAGCCCTTCTTATCGCAATGATGGCAATTTGT 7020  
Db 6961 CACTGTGCGTAGAGGCACTCTTGAACGATAGCCCTTCTTATCGCAATGATGGCAATTTGT 7020  
QY 7021 AGAAGCCATCTTCTTCTACTGTCTTTCGATGAAGTGACAGATAGCTGGGCAATGGA 7080  
Db 7021 AGAAGCCATCTTCTTCTACTGTCTTTCGATGAAGTGACAGATAGCTGGGCAATGGA 7080  
QY 7081 ATCCGAGAGGTTTCCGATATATACCTTGTGAAAGTCTCAATAGCCCTCTGGTCTT 7140  
Db 7081 ATCCGAGAGGTTTCCGATATATACCTTGTGAAAGTCTCAATAGCCCTCTGGTCTT 7140  
QY 7141 CTGAGACTGTATCTTTTGATATTTTGGAGTAGACGAGAGTGTGCTCCACCATGTTGG 7200  
Db 7141 CTGAGACTGTATCTTTTGATATTTTGGAGTAGACGAGAGTGTGCTCCACCATGTTGG 7200  
QY 7201 GGACTCTAGATATACATCAATCCACTTGTCTTTGAAAGAGTGGTTGGAAGCTTCTCTTTT 7260  
Db 7201 GGACTCTAGATATACATCAATCCACTTGTCTTTGAAAGAGTGGTTGGAAGCTTCTCTTTT 7260  
QY 7261 CCACATGTTCTGCTGGTGGGGTCCATCTTTTGGGACCACTGTGCGTAGAGGCACTTT 7320  
Db 7261 CCACATGTTCTGCTGGTGGGGTCCATCTTTTGGGACCACTGTGCGTAGAGGCACTTT 7320  
QY 7321 GAAACGATAGCCTTCTCTTTATCGCAATGATGCAATTTGTAGAAGCCATCTCTCTTTCTA 7380  
Db 7321 GAAACGATAGCCTTCTCTTTATCGCAATGATGCAATTTGTAGAAGCCATCTCTCTTTCTA 7380  
QY 7381 CTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATA 7440  
Db 7381 CTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATA 7440  
QY 7441 TTACCCCTTTGTGAAAGTCTCAATAGGCCCTCTGCTCTTGAGACTGTATCTTTGAT 7500

Db 7441 TTACCCCTTTGTGAAAGTCTCAATAGGCCCTCTGCTCTTGAGACTGTATCTTTGATAT 7500  
QY 7501 TCTTGGAGTAGACGAGAGTGTGCTGTCCACCAATGTTGGGATCCACTAGTTCTAGAGCG 7560  
Db 7501 TCTTGGAGTAGACGAGAGTGTGCTGTCCACCAATGTTGGGATCCACTAGTTCTAGAGCG 7560  
QY 7561 GCCGCCACCGCGGTGGAGCT 7580  
Db 7561 GCCGCCACCGCGGTGGAGCT 7580  
RESULT 2  
ADL71846  
ID ADL71846 standard; DNA; 1746 BP.  
XX ADL71846;  
AC ADL71846;  
XX 20-MAY-2004 (first entry)  
XX Arabidopsis thaliana OBP3 antisense DNA #7.  
XX Transgenic plant; Dof transcription factor; ocs binding factor;  
KW plant size; plant stature; root growth; plant; gene; ds; OBF;  
KW OBF binding protein; OBP3; SOBI; mouse-ear cross.  
XX Arabidopsis thaliana.  
XX US2004045055-A1.  
XX 04-MAR-2004.  
XX 28-AUG-2003; 2003US-00650249.  
XX 28-AUG-2002; 2002US-0406657P.  
XX: (UNIW ) UNIV WASHINGTON.  
PA Neff MM;  
XX WPI; 2004-225757/21.  
XX New transgenic plant cell, useful in producing plants with altered size  
PT and stature and with normal and healthy root growth.  
PS Claim 38; SEQ ID NO 18; 53pp; English.  
XX The invention relates to a transgenic plant transformed by a Dof  
CC transcription factor, OBF (ocs binding factor) binding protein (OBP3).  
CC OBP3 is also known as SOBI. The transgenic plant cell and OBP3 nucleic  
CC acid and polypeptides are useful in producing transgenic plants with  
CC altered size and stature and with normal and healthy root growth. The  
CC present sequence is Arabidopsis thaliana OBP3 antisense DNA.  
XX Sequence 1746 BP; 560 A; 304 C; 304 G; 578 T; 0 U; 0 Other;  
SQ  
Query Match 22.8%; Score 1730; DB 12; Length 1746;  
Best Local Similarity 99.8%; Pred. No. 2.6e-258;  
Matches 1743; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 4436 AACAGAGAAGCAAAATCCAGATCGAAATCTAGCGTGGTCTCGACTGATAATACTACT 4495  
Db 1 AACAGAGAAGCAAAATCCAGATCGAAATCTAGCGTGGTCTCGACTGATAATACTACT 60  
QY 4496 AGTACTTCACTTACTTCTCGCCCAAGTTACTCAAAACCTTAGCAAGTTTTCATAGCTAC 4555  
Db 61 AGTACTTCACTTACTTCTCGCCCAAGTTACTCAAAACCTTAGCAAGTTTTCATAGCTAC 120  
QY 4556 GGTCAAAATCCCGAGTTTAAATTCCAACTTGCCCACTTTGCTCTCTCCAAAGCCTTGA 4615  
Db 121 GGTCAAAATCCCGAGTTTAAATTCCAACTTGCCCACTTTGCTCTCTCCAAAGCCTTGA 180  
QY 4616 GATTACAATTCAGCAACACTGGATTAGATTTTGGTGGAACTCAAAATAGCAACATGATA 4675





CC in a plant. The vector of the invention is useful for analyzing a plant  
CC gene. The present DNA sequence is claimed in the invention.  
XX  
SQ Sequence 17511 BP; 4453 A; 4201 C; 4511 G; 4344 T; 0 U; 2 Other;  
Query Match 18.0%; Score 1367.4; DB 14; Length 17511;  
Best Local Similarity 99.9%; Pred. No. 3.4e-202;  
Matches 1368; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 6177 TTGTTTCCGATCTAGATATCATCAATCCACTTCTTTGAAGACGTGGTGGAAAGCTCT 6236  
DB 17505 TAGTTTCGATCTAGATATCATCAATCCACTTCTTTGAAGACGTGGTGGAAAGCTCT 17446  
QY 6237 TCTTTTTCACGATGTTCTCGTGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAG 6296  
DB 17445 TCTTTTTCACGATGTTCTCGTGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAG 17386  
QY 6297 GCATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCATTTCTGAGAAGCATCTTCC 6356  
DB 17385 GCATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCATTTCTGAGAAGCATCTTCC 17326  
QY 6357 TTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTT 6416  
DB 17325 TTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTT 17266  
QY 6417 CCGGATATACCTTTTGTGAAAAGTCTCAATAGCCCTCTGTCCTTCTGAGACTGTATCT 6476  
DB 17265 CCGGATATATACCTTTTGTGAAAAGTCTCAATAGCCCTCTGTCCTTCTGAGACTGTATCT 17206  
QY 6477 TTGATATCTTCGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCA 6536  
DB 17205 TTGATATCTTCGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCA 17146  
QY 6537 CATCAATCCACTTGTCTTGAACACGTGGTGGAAAGTCTTCTTTTCCACGATGTTCTCT 6596  
DB 17145 CATCAATCCACTTGTCTTGAACACGTGGTGGAAAGTCTTCTTTTCCACGATGTTCTCT 17086  
QY 6597 GTGGTGGGGTCCATCTTTTGGGACCACTGTGCGTAGAGGCATCTTGAACGATAGCCTTT 6656  
DB 17085 GTGGTGGGGTCCATCTTTTGGGACCACTGTGCGTAGAGGCATCTTGAACGATAGCCTTT 17026  
QY 6657 CTTTATCCGATGATGGCATTTGTAGAGCCATCTTCTTCTACTGTCCTTTCGATG 6716  
DB 17025 CTTTATCCGATGATGGCATTTGTAGAGCCATCTTCTTCTACTGTCCTTTCGATG 16966  
QY 6717 AAGTCACATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATATACCCCTTTGTGA 6776  
DB 16965 AAGTCACATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATATACCCCTTTGTGA 16906  
QY 6777 AAGTCACATAGCCTCTGTCCTTCTGAGACTGTATCTTTGATATCTTGGAGTAGAGC 6836  
DB 16905 AAGTCACATAGCCTCTGTCCTTCTGAGACTGTATCTTTGATATCTTGGAGTAGAGC 16846  
QY 6837 AGAGTGTCTGTCACCAATGTTGGGATCTAGATATACATCAATCCACTTGTCTTGA 6896  
DB 16845 AGAGTGTCTGTCACCAATGTTGGGATCTAGATATACATCAATCCACTTGTCTTGA 16786  
QY 6897 GAGCTGTTTGAACGCTTCTTTTCCAGATGTTCTCTGTCGGTGGGGTCCATCTTTG 6956  
DB 16785 GAGCTGTTTGAACGCTTCTTTTCCAGATGTTCTCTGTCGGTGGGGTCCATCTTTG 16726  
QY 6957 GGAACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCAT 7016  
DB 16725 GGAACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCAT 16665  
QY 7017 TTGTAGAAGCATCTTCTTTTCTACTGTCCTTTCCATGAGTACAGATAGCTGGGCA 7076  
DB 16665 TTGTAGAAGCATCTTCTTTTCTACTGTCCTTTCCATGAGTACAGATAGCTGGGCA 16606  
QY 7077 TGAATCCGAGGAGTTTCCGATATATACCTTTGTTGAAAAGTCTCAATAGCCCTCTCG 7136  
DB 16605 TGAATCCGAGGAGTTTCCGATATATACCTTTGTTGAAAAGTCTCAATAGCCCTCTCG 16546

QY 7137 TCTTCTGAGACTGTATCTTTGATATCTTGGAGTAGACGAGAGTGTGTCCTCCACCATG 7196  
DB 16545 TCTTCTGAGACTGTATCTTTGATATCTTGGAGTAGACGAGAGTGTGTCCTCCACCATG 16486  
QY 7197 TTGGGATCTAGATATACATCAATCCACTTCTTTGAAGACGTGGTGGAAAGCTCTTCT 7256  
DB 16485 TTGGGATCTAGATATACATCAATCCACTTCTTTGAAGACGTGGTGGAAAGCTCTTCT 16426  
QY 7257 TTTTCCACGATGTTCTCTGTCGGTGGGGTCCATCTTTGGGACCACTGTCCGTAGAGCA 7316  
DB 16425 TTTTCCACGATGTTCTCTGTCGGTGGGGTCCATCTTTGGGACCACTGTCCGTAGAGCA 16366  
QY 7317 TCTTGAACGATAGCCTTTTCTTTATCGCAATGATGGCATTTGTAGAAGCATCTTCTTT 7376  
DB 16365 TCTTGAACGATAGCCTTTTCTTTATCGCAATGATGGCATTTGTAGAAGCATCTTCTTT 16306  
QY 7377 TCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCC 7436  
DB 16305 TCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCC 16246  
QY 7437 GATATTACCTTTTGTGAAAAGTCTCAATAGCCCTCTGTCCTTCTGAGACTGTATCTTTG 7496  
DB 16245 GATATTACCTTTTGTGAAAAGTCTCAATAGCCCTCTGTCCTTCTGAGACTGTATCTTTG 16186  
QY 7497 ATATTCTTGGAGTAGACGAGAGTGTGTCCTCCACCATGTTGGGATCC 7545  
DB 16185 ATATTCTTGGAGTAGACGAGAGTGTGTCCTCCACCATGTTGGGATCC 16137

RESULT 5

ADV39062/c  
ID ADV39062 standard; DNA; 18987 BP.  
XX AC ADV39062;  
XX DT 24-FEB-2005 (first entry)  
XX DE Plant gene activation vector-related DNA sequence pHR-AT-GFP - SEQ ID 14.  
XX KW gene activation vector; genetic analysis; ds.  
XX OS Unidentified.  
XX PN WO2004106520-A1.  
XX PD 09-DEC-2004.  
XX PF 28-MAY-2004; 2004WO-JP007789.  
XX PR 30-MAY-2003; 2003JP-00153985.  
XX PA (RIKE ) RIKEN KK.  
XX PI Muranaka T, Seki H;  
XX DR WPI; 2005-021291/02.  
XX PT Novel plant gene active vector comprising hairy-root induction gene, and  
PT enhancer and/or promoter that functions in plant, useful for activating  
PT plant gene forming hairy root.  
XX PS Claim 5; SEQ ID NO 14; 54pp; Japanese.  
XX CC The invention comprises a plant gene activation vector that contains a  
CC hairy-root induction gene and enhancer and/or promoter which can function  
CC in a plant. The vector of the invention is useful for analyzing a plant  
CC gene. The present DNA sequence is claimed in the invention.  
XX SQ Sequence 18987 BP; 4871 A; 4609 C; 4863 G; 4642 T; 0 U; 2 Other;

Query Match 18.0%; Score 1367.4; DB 14; Length 18987;  
Best Local Similarity 99.9%; Pred. No. 3.4e-202;  
Matches 1368; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	6177	TTGTTTCGATCTAGATATCAATCAATCCACTTGTCTTTGAAGACGTGTT	6236	QY	7257	TTTTCCACGATGTTCTCGTGGGTGGGTCCATCTTTGGGACCACCTGTCGGTAGAGCA	7316
Db	18981	TAGTTTCGATCTAGATATCAATCAATCCACTTGTCTTTGAAGACGTGTT	18922	Db	17901	TTTTCCACGATGTTCTCGTGGGTGGGTCCATCTTTGGGACCACCTGTCGGTAGAGCA	17842
QY	6237	TCCTTTTCCACGATGTTCTCGTGGGTGGGTCCATCTTTGGGACCACCTGTCGGTAGAG	6296	QY	7317	TCCTGAAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTCTTT	7376
Db	18921	TCCTTTTCCACGATGTTCTCGTGGGTGGGTCCATCTTTGGGACCACCTGTCGGTAGAG	18862	Db	17841	TCCTGAAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTCTTT	17782
QY	6297	GCATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTC	6356	QY	7377	TCCTGATCTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC	7436
Db	18861	GCATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTC	18802	Db	17781	TCCTGATCTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC	17722
QY	6357	TTTTCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTT	6416	QY	7437	GATATTAACCTTTGTTGAAAAGTCTCAATAGCCCTCTCTGAGACTGATCTTTG	7496
Db	18801	TTTTCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTT	18742	Db	17721	GATATTAACCTTTGTTGAAAAGTCTCAATAGCCCTCTCTGAGACTGATCTTTG	17662
QY	6417	CCCGATATTACCTTTGTTGAAAAGTCTCAATAGCCCTCTCTGAGACTGATCTT	6476	QY	7497	ATATTCTTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGGATCC	7545
Db	18741	CCCGATATTACCTTTGTTGAAAAGTCTCAATAGCCCTCTCTGAGACTGATCTT	18682	Db	17661	ATATTCTTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGGATCC	17613
QY	6477	TTGATATTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGGATCTAGATATCA	6536	RESULT 6			
Db	18681	TTGATATTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGGATCTAGATATCA	18622	AAA88401/c			
QY	6537	CATCAATCCACTTGTCTTGAAGACGTGTTGGAACGTCTCTTTTCCACGATGTTCTC	6596	ID AAA88401 standard; DNA; 1361 BP.			
Db	18621	CATCAATCCACTTGTCTTGAAGACGTGTTGGAACGTCTCTTTTCCACGATGTTCTC	18562	XX AAA88401;			
QY	6597	GTGGGTGGGGTCCATCTTTGGGACCACCTGTCGGTAGAGGCATCTTGAACGATAGCCTTT	6656	XX 09-JAN-2001 (first entry)			
Db	18561	GTGGGTGGGGTCCATCTTTGGGACCACCTGTCGGTAGAGGCATCTTGAACGATAGCCTTT	18502	XX 4X CaMV 35S enhancer construct.			
QY	6657	CTTTATCCCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGCTCTTCGATG	6716	CaMV; 35S; enhancer; trait-associated gene identification;			
Db	18501	CTTTATCCCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGCTCTTCGATG	18442	activation tagging; fruit; transgenic plant; ss.			
QY	6717	AAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTGTGA	6776	Cauliflower mosaic virus.			
Db	18441	AAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTGTGA	18382	Key Location/Qualifiers			
QY	6777	AAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTGTGA	6836	repeat_region 1..1354			
Db	18381	AAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTGTGA	18322	/tag= a			
QY	6837	AGAGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCAATCAATCAATCAAT	6896	/repeat_type= TANDEM			
Db	18321	AGAGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCAATCAATCAATCAAT	18262	/note= "4 CaMV 35S enhancer units"			
QY	6897	GACGTGGTTGGAACGTCTTCTTTTCCACGATGTTCTCGTGGGTGGGTCCATCTTTG	6956	repeat_unit 1..339			
Db	18261	GACGTGGTTGGAACGTCTTCTTTTCCACGATGTTCTCGTGGGTGGGTCCATCTTTG	18202	/tag= b			
QY	6957	GGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCA	7016	/note= "CaMV 35S enhancer unit 1"			
Db	18201	GGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCA	18142	/tag= c			
QY	7017	TTGTAGAAGCCATCTTCTTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGCA	7076	/note= "CaMV 35S enhancer AluI-EcoRV fragment"			
Db	18141	TTGTAGAAGCCATCTTCTTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGCA	18082	/tag= d			
QY	7077	TGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGAAAAGTCTCAATAGCCCTCTGG	7136	/note= "129 bp fragment of the CaMV sequence"			
Db	18081	TGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGAAAAGTCTCAATAGCCCTCTGG	18022	/tag= e			
QY	7137	TCCTTCGAGACTGTATCTTTGATATTCTTTGGAGTAGACGAGAGTGTGTCCTCCACCATG	7196	/note= "additional 7 bp not associated with 35S enhancer"			
Db	18021	TCCTTCGAGACTGTATCTTTGATATTCTTTGGAGTAGACGAGAGTGTGTCCTCCACCATG	17962	/tag= f			
QY	7197	TTGGGATCTAGATATCAATCAATCCACTTGTCTTTGAAGACGTGTTGGAACGTCTTCT	7256	/note= "CaMV 35S enhancer unit 2"			
Db	17961	TTGGGATCTAGATATCAATCAATCCACTTGTCTTTGAAGACGTGTTGGAACGTCTTCT	17902	/tag= g			
				/note= "CaMV 35S enhancer AluI-EcoRV fragment"			
				/tag= h			
				/note= "129 bp fragment of the CaMV sequence"			
				/tag= i			
				/note= "additional 7 bp not associated with 35S enhancer"			
				/tag= j			
				/note= "CaMV 35S enhancer unit 3"			
				/tag= k			
				/note= "CaMV 35S enhancer AluI-EcoRV fragment"			
				/tag= l			

```
FT enhancer /note= "129 bp fragment of the CamV sequence" 1011. .1017 Db 1300 TTTTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGC 1241
FT repeat_unit /tag= m /note= "additional 7 bp not associated with 35S enhancer" 1018. .1354 QY 6299 ATCTTGAACGATAGCTTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCCTT 6358
FT enhancer /tag= a /note= "CamV 35S enhancer unit 4" 1018. .1219 Db 1240 ATCTTGAACGATAGCTTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCCTT 1181
FT enhancer /tag= n /note= "CamV 35S enhancer AluI-EcoRV fragment" 1018. .1348 QY 6359 TTCTACTGTCCTTTCCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTC 6418
FT enhancer /tag= o /note= "129 bp fragment of the CamV sequence" 1018. .1354 Db 1180 TTCTACTGTCCTTTCCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTC 1121
FT enhancer /tag= p /note= "additional 6 bp not associated with 35S enhancer" 1018. .1354 QY 6419 CGATATTACCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACCTGATCTTT 6478
FT WO200053794-A2. /note= "additional 6 bp not associated with 35S enhancer" 1018. .1354 Db 1120 CGATATTACCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACCTGATCTTT 1061
FT 14-SEP-2000. QY 6479 GATATTCTTGAGTAGACAGAGTCTGCTGCTCCACCATGTTGGGATCTAGATATCACA 6538
FT 09-MAR-2000; 2000WO-US006298. Db 1060 GATATTCTTGAGTAGACAGAGTCTGCTGCTCCACCATGTTGGGATCTAGATATCACA 1001
FT 12-MAR-1999; 99US-0124232P. QY 6539 TCAATCCACTTGTCTTTGAAAGACGTGTTGGAACGTCTCTTTTCCACGATGTTCTCTGT 6598
FT (AGRI-) AGRITOPE INC. Db 1000 TCAATCCACTTGTCTTTGAAAGACGTGTTGGAACGTCTCTTTTCCACGATGTTCTCTGT 941
FT Wagner R, Mathews H, Liu XL, Waggoner WJ; QY 6599 GGGTGGGGTCCATCTTTGGGACCACTGCTCGTAGAGGACCTCTTGAACGATAGCCCTTCC 6658
FT WPI; 2000-594336/56. Db 940 GGGTGGGGTCCATCTTTGGGACCACTGCTCGTAGAGGACCTCTTGAACGATAGCCCTTCC 881
FT Identifying genes associated with a desired trait for isolating and QY 6659 TTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTCTACTGCTCTTTCGATGAA 6718
FT Characterizing the genes comprises using an enhancer element which Db 880 TTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTCTACTGCTCTTTCGATGAA 821
FT enhances gene expression and stably integrates into the plant genome. QY 6719 GTGACAGATAGCTGGGCAATCGAATCCGAGAGGTTTCCCGATATTAACCTTTGTTGAAA 6778
FT Claim 4; Fig 5; 42pp; English. Db 820 GTGACAGATAGCTGGGCAATCGAATCCGAGAGGTTTCCCGATATTAACCTTTGTTGAAA 761
FT The present sequence is that of a 4X cauliflower mosaic virus (CaMV) 35S QY 6779 AGTCTCAATAGCCCTCTGCTCTTTGAGACCTGATCTTTGATATTTCTTGGAGTAGACGAG 6838
FT enhancer sequence preferred for use in the method of the invention. It Db 760 AGTCTCAATAGCCCTCTGCTCTTTGAGACCTGATCTTTGATATTTCTTGGAGTAGACGAG 701
FT includes 4 repeats of 202 bp AluI-EcoRV fragments of the 35S enhancer, QY 6839 AGTCTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCAATCTTCTTTGAGA 6898
FT 129 bp of the CaMV sequence associated with each tandem Alu-EcoRV repeat, Db 700 AGTCTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCAATCTTCTTTGAGA 641
FT and an additional 7 bp repeated sequence, which does not appear in the QY 6899 CGTGTGTGGAAGCTCTTCTTTTCCAGATGTTCTCTGCTGGGTGGGGTCCATCTTTGGG 6958
FT 35S enhancer region of the native CaMV genome. This 4X CaMV 35S enhancer Db 640 CGTGTGTGGAAGCTCTTCTTTTCCAGATGTTCTCTGCTGGGTGGGGTCCATCTTTGGG 581
FT element can be used in a method for identifying genes associated with a QY 6959 ACCACTGTGCTGAGGACATCTTGAACGATAGCCCTTTCTTTATCGCAATGATGGCATTT 7018
FT desired trait in a fruit-bearing plant. The method involves: transforming Db 580 ACCACTGTGCTGAGGACATCTTGAACGATAGCCCTTTCTTTATCGCAATGATGGCATTT 521
FT plant cells with an activation tagging vector comprising an element which QY 7019 GTAGAAGCCATCTTCTTTCTACTGCTCTTTCGATGAAGTCAACAGATAGCTGGCAATG 7078
FT functions to enhance gene expression and has the ability to integrate Db 520 GTAGAAGCCATCTTCTTTCTACTGCTCTTTCGATGAAGTCAACAGATAGCTGGCAATG 461
FT into the plant genome in a manner effective to enhance expression of QY 7079 GAATCCGAGGAGTTCCTCGATATTAACCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTG 7138
FT native plant genes, selecting transformed plant cells, regenerating Db 460 GAATCCGAGGAGTTCCTCGATATTAACCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTG 401
FT transformed plant cells to yield mature plants, selecting plants having a QY 7139 TTTCTGAGACTGATCTTTTGATATTTCTTGAGTAGACGAGAGTGTGCTGTCCACCAATGTT 7198
FT desired trait, identifying, isolating and characterizing genes the Db 400 TTTCTGAGACTGATCTTTTGATATTTCTTGAGTAGACGAGAGTGTGCTGTCCACCAATGTT 341
FT transcription of which has been enhanced, and confirming the contribution QY 7199 GGGGATCTAGATATCACATCAATCCATTCCTTTGAGACGCTGTTGGAACGCTCTCTTTT 7258
FT of the modified expression of each identified gene to the desired trait. Db 340 GGGGATCTAGATATCACATCAATCCATTCCTTTGAGACGCTGTTGGAACGCTCTCTTTT 281
FT The desired trait may be increased resistance to fungal, bacterial or QY 7259 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATC 7318
FT viral pathogens, insects, modifications in flower size, flower number, Db 280 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATC 221
FT flower pigmentation and shape, modified leaf number, leaf pigmentation QY 7319 TTGAAACGATAGCTTTCTTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTCTTTTC 7378
FT and shape, modified seed number, pattern or distribution of leaves and
FT flowers, modified stem length or increased nodes, root mass or root
FT development characteristics or increased drought, salt and antibiotic
FT tolerance. Plants having short life cycles are transformed, as
FT exemplified by dwarf varieties of tomato
SQ Query Match 17.9%; Score 1360; DB 3; Length 1361;
Best Local Similarity 100.0%; Pred. No. 4.2e-201;
Matches 1360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6179 GTTTCGGATCTAGATATCACATCAATCCACTTCTTTGAGAGCGTGGTGGAAACGCTCTTC 6238
Db 1360 GTTTCGGATCTAGATATCACATCAATCCACTTCTTTGAGAGCGTGGTGGAAACGCTCTTC 1301
QY 6239 TTTTCCACGATGTTCTCTGCTGGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGC 6298
```



Db 220 TTGAAGATAGCTTCTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTCTTTC 161  
QY 7379 TACTGTCTCTTTCGATGAAGTGACAGATAGCTGGGCAATGGAAATCCGAGAGGTTCCTCCGA 7438  
Db 160 TACTGTCTCTTTCGATGAAGTGACAGATAGCTGGGCAATGGAAATCCGAGAGGTTCCTCCGA 101  
QY 7439 TATTACCTCTTTCGATGAAGTGACAGATAGCTGGGCAATGGAAATCCGAGAGGTTCCTCCGA 7498  
Db 100 TATTACCTCTTTCGATGAAGTGACAGATAGCTGGGCAATGGAAATCCGAGAGGTTCCTCCGA 41  
QY 7499 ATTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTG 7538  
Db 40 ATTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTG 1

RESULT 7  
ADL71845  
ID ADL71845 standard; DNA; 1235 BP.  
XX  
AC ADL71845;  
XX  
20-MAY-2004 (first entry)  
XX  
DE Arabidopsis thaliana OBP3 antisense DNA #6.  
XX  
KW Transgenic plant; Dof transcription factor; ocs binding factor;  
KW plant size; plant stature; root growth; plant; gene; ds; OBF;  
KW OBF binding protein; OBP3; SOB1; SOB1; mouse-ear cress.  
XX  
OS Arabidopsis thaliana.  
XX  
PN US2004045055-A1.  
XX  
PD 04-MAR-2004.  
XX  
PF 28-AUG-2003; 2003US-00650249.  
XX  
PR 28-AUG-2002; 2002US-0406657P.  
XX  
PA (UNIW ) UNIV WASHINGTON.  
XX  
PI Neff MM;  
XX  
WPI; 2004-225757/21.  
XX  
New transgenic plant cell, useful in producing plants with altered size  
and stature and with normal and healthy root growth.  
XX  
Claim 37; SEQ ID NO 17; 53pp; English.  
XX  
The invention relates to a transgenic plant transformed by a Dof  
transcription factor, OBF (ocs binding factor) binding protein (OBP3).  
CC OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic  
CC acid and polypeptides are useful in producing transgenic plants with  
CC altered size and stature and with normal and healthy root growth. The  
CC present sequence is Arabidopsis thaliana OBP3 antisense DNA.  
XX  
SQ Sequence 1235 BP; 408 A; 222 C; 179 G; 426 T; 0 U; 0 Other;  
Query Match 16.3%; Score 1235; DB 12; Length 1235;  
Best Local Similarity 100.0%; Pred. No. 9e-182;  
Matches 1235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3045 CAACCAAGACGATGCGTATATGATGCTTCCAAAAGGTTAGTAAGTAAGAGGCTTTTATTTATGAAA 3104  
Db 1 CAACCAAGACGATGCGTATATGATGCTTCCAAAAGGTTAGTAAGTAAGAGGCTTTTATTTATGAAA 60  
QY 3105 CAATCGACACTTAATTCGAAAAGGTTAGTAAGTAAGAGGCTTTTATTTATGAAA 3164  
Db 61 CAATCGACACTTAATTCGAAAAGGTTAGTAAGTAAGAGGCTTTTATTTATGAAA 120  
QY 3165 ACAAAGAAATAAAGAGGCTTAAGAAATGATGAAATTTGAAAGAGAAAAAGAGCATTTG 3224

Db 121 ACAAAGAAATAAAGAGGCTTAAGAAATGATGAAATTTGAAAGAGAAAAAGAGCATTTG 180  
QY 3225 TTATAGAAAAAGAAAAAGAGAGAGTAAAGAGAAATTAAGAAAAACAATAAATTAAGCAA 3284  
Db 181 TTATAGAAAAAGAAAAAGAGAGAGTAAAGAGAAATTAAGAAAAACAATAAATTAAGCAA 240  
QY 3285 AGGAAACCTTCAATTTCTCTTATCCCATTCAGCTCTCCCTCTCTCTCTCTCTCTCTCTCT 3344  
Db 241 AGGAAACCTTCAATTTCTCTTATCCCATTCAGCTCTCCCTCTCTCTCTCTCTCTCTCT 300  
QY 3345 CTCTCTCTCTAGATCAATTTCTTCTTATGATGTGATTAATCAACCATATCTCGGACC 3404  
Db 301 CTCTCTCTCTAGATCAATTTCTTCTTATGATGTGATTAATCAACCATATCTCGGACC 360  
QY 3405 TCTTACCTAAAAAGGATACAAAGATCAAAAGATGGTTTTCTCATCTCTCCAGTG 3464  
Db 361 TCTTACCTAAAAAGGATACAAAGATCAAAAGATGGTTTTCTCATCTCTCCAGTG 420  
QY 3465 AATCAGTTTCGATTCCTCAAAATTTGGCAGCAGGTAAAAAATCAGTTTATGATATTTGCTAGAT 3524  
Db 421 AATCAGTTTCGATTCCTCAAAATTTGGCAGCAGGTAAAAAATCAGTTTATGATATTTGCTAGAT 480  
QY 3525 GTTCTGATTCGTTCTCTTCTTCTCAAGCTCGATCAAGATTTATGAAAAATTTGATGAGA 3584  
Db 481 GTTCTGATTCGTTCTCTTCTTCTCAAGCTCGATCAAGATTTATGAAAAATTTGATGAGA 540  
QY 3585 TTTTGTTCGACAAATTCCTAGCTTATGTGACGGCGATATATATCTTATGATATTTGCTAGAT 3644  
Db 541 TTTTGTTCGACAAATTCCTAGCTTATGTGACGGCGATATATATCTTATGATATTTGCTAGAT 600  
QY 3645 TTAGTTGATTAACCCCTTTTCTTCTTCTTCTCGAATATACGAAATATATAAAGAT 3704  
Db 601 TTAGTTGATTAACCCCTTTTCTTCTTCTTCTCGAATATACGAAATATATAAAGAT 560  
QY 3705 GATTTCAATTTTGGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3764  
Db 661 GATTTCAATTTTGGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720  
QY 3765 TAAAAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3824  
Db 721 TAAAAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780  
QY 3825 AATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3884  
Db 781 AATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840  
QY 3885 GTTTAAAGTTACTATGACTTTTAACTGAGTTTATTTATCCATTTCTTTTTCGAGCTTTCT 3944  
Db 841 GTTTAAAGTTACTATGACTTTTAACTGAGTTTATTTATCCATTTCTTTTTCGAGCTTTCT 900  
QY 3945 TGAAAAACTATAATTAATCTGCAATTTCTGTCAAAGTAGTCACAAATTTTATCTATTTTC 4004  
Db 901 TGAAAAACTATAATTAATCTGCAATTTCTGTCAAAGTAGTCACAAATTTTATCTATTTTC 960  
QY 4005 TTTTGTCTCCGACCAATGTTTCAAACTCGAATCCTTTGGTTAAAGTTGTTTCTGCTTTA 4064  
Db 961 TTTTGTCTCCGACCAATGTTTCAAACTCGAATCCTTTGGTTAAAGTTGTTTCTGCTTTA 1020  
QY 4065 TTATAAACCCTGAACTAAATTAAGTACAAATTTATGTAATGAGCAAGGAAACCAACATC 4124  
Db 1021 TTATAAACCCTGAACTAAATTAAGTACAAATTTATGTAATGAGCAAGGAAACCAACATC 1080  
QY 4125 AGCTAGAATGTGTCAACAATGACCAAGACCCCTTAATTAATTAATTAATTAATTAATTAAT 4184  
Db 1081 AGCTAGAATGTGTCAACAATGACCAAGACCCCTTAATTAATTAATTAATTAATTAATTAAT 1140  
QY 4185 CACCGACTTCTCAGGTTGAGGTTTCGAGTCAAGCTAGAGTGAATCAATGGTGGACG 4244  
Db 1141 CACCGACTTCTCAGGTTGAGGTTTCGAGTCAAGCTAGAGTGAATCAATGGTGGACG 1200  
QY 4245 CTCGGATCGCAAAAGTCCCATTTGCTGAGGAGCT 4279  
Db 1201 CTCGGATCGCAAAAGTCCCATTTGCTGAGGAGCT 1235

```
RESULT 8
ADL71848
ID ADL71848 standard; DNA; 1062 BP.
XX AC ADL71848;
XX DT 20-MAY-2004 (first entry)
XX DE Arabidopsis thaliana OBP3 antisense DNA #9.
XX KW Transgenic plant; Dof transcription factor; ocs binding factor;
XX KW plant size; plant stature; root growth; plant; gene; ds; OBF;
XX KW OBF binding protein; OBP3; SOB1; mouse-ear cress.
XX OS Arabidopsis thaliana.
XX PN US2004045055-A1.
XX PD 04-MAR-2004.
XX PF 28-AUG-2003; 2003US-00650249.
XX PR 28-AUG-2002; 2002US-0406657P.
XX PA (UNITW ) UNIV WASHINGTON.
XX PI Neff MM;
XX DR WPI; 2004-225757/21.
XX PT New transgenic plant cell, useful in producing plants with altered size
XX PT and stature and with normal and healthy root growth.
XX PS Claim 40; SEQ ID NO 20; 53bp; English.
XX CC The invention relates to a transgenic plant transformed by a Dof
XX CC transcription factor, OBF (ocs binding factor) binding protein (OBP3).
XX CC OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic
XX CC acid and polypeptides are useful in producing transgenic plants with
XX CC altered size and stature and with normal and healthy root growth. The
XX CC present sequence is Arabidopsis thaliana OBP3 antisense DNA.
XX SQ Sequence 1062 BP; 348 A; 159 C; 155 G; 400 T; 0 U; 0 Other;

Query Match 14.0%; Score 1062; DB 12; Length 1062;
Best Local Similarity 100.0%; Pred. No. 5,1e-155;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5117 TAAGTACTCAGCACTAGCTATCTTGATGATCTCTTTGTTGGTGGGGTGATCATTTGGTG 5176
DB 1 TAAGTACTCAGCACTAGCTATCTTGATGATCTCTTTGTTGGTGGGGTGATCATTTGGTG 60
QY 5177 CTGTGTCATCGAGTATTCTGCTGAGGAGATCAACCATGCGCTATATCCAAAGGCTAAT 5236
DB 61 CTGTGTCATCGAGTATTCTGCTGAGGAGATCAACCATGCGCTATATCCAAAGGCTAAT 120
QY 5237 TTTTGAGGCTCAAAAGGAGGTATGGTTATAAAACTATCTTTTGTGATCTTTTAAAGATCT 5296
DB 121 TTTTGAGGCTCAAAAGGAGGTATGGTTATAAAACTATCTTTTGTGATCTTTTAAAGATCT 180
QY 5297 TCAAGTGTGAGTATGTTTATTTGGTGGCTTCTGGTGATATTATGTTTATAGAAATTT 5356
DB 181 TCAAGTGTGAGTATGTTTATTTGGTGGCTTCTGGTGATATTATGTTTATAGAAATTT 240
QY 5357 GGTCTTATATATGGCTATATAGAGTGTGGGTGATATGATGATGATGATGATGATGAT 5416
DB 241 GGTCTTATATATGGCTATATAGAGGTTGGGTGATATGATGATGATGATGATGATGAT 300
QY 5417 GTTGGAAACTTTTCTGTGTGTTCAATGAATATCATCGAAATCTCAATTTCTTTGGAGACC 5476
DB 301 GTTGGAACTTTTGTGTGTTCAATGAATATCATCGAAATCTCAATTTCTTTGGAGACC 360
```

```
QY 5477 CATTATGAGACATTGAGACATCTATAGAACATATATGTAATGTAATATAAAGCTACTTAA 5536
DB 361 CATTATGAGACATTGAGACATCTATAGAACATATATGTAATGTAATATAAAGCTACTTAA 420
QY 5537 GTCGAAATTTTATGACCAAAGTAAATTAATGCGAATGTACATGCTAAATATCGAGTTT 5596
DB 421 GTCGAAATTTTATGACCAAAGTAAATTAATGCGAATGTACATGCTAAATATCGAGTTT 480
QY 5597 AAACATATTTTTCGAATATACCACTATTTCTCTTTGTCCTCAACTATATATATCTTAT 5656
DB 481 AAACATATTTTTCGAATATACCACTATTTCTCTTTGTCCTCAACTATATATATCTTAT 540
QY 5657 CTGATTTCTTATTTTCTTTTAAATTCCTTTTCTTTTCCCAAGACACAAAAA 5716
DB 541 CTGATTTCTTATTTTCTTTTAAATTCCTTTTCTTTTCCCAAGACACAAAAA 600
QY 5717 AATCAGAAACGAAAAAAGAGATTTTAAATTCATAACCCACGAGAAATATGACCTTA 5776
DB 601 AATCAGAAACGAAAAAAGAGATTTTAAATTCATAACCCACGAGAAATATGACCTTA 660
QY 5777 AATTCAGACTAATCCCCCAAAATTTTCAGAAATTTATGTAATTTTGGGATTTAATAT 5836
DB 661 AATTCAGACTAATCCCCCAAAATTTTCAGAAATTTATGTAATTTTGGGATTTAATAT 720
QY 5837 TCACAAATCATATGCGCAACTAACTAATTTGAAAGACAAATGGAATGAAACCATGCA 5896
DB 721 TCACAAATCATATGCGCAACTAATTTGAAAGACAAATGGAATGAAACCATGCA 780
QY 5897 TAAATCTCTCAAGTCTCAACCTATGAAAGATCATGTAAACCAATAGACTATCATGATTA 5956
DB 781 TAAATCTCTCAAGTCTCAACCTATGAAAGATCATGTAAACCAATAGACTATCATGATTA 840
QY 5957 GTTAATGATGATCTATATGTAATTTCTTTGAAACATAGATATGTAATTTCTGGATATA 6016
DB 841 GTTAATGATGATCTATATGTAATTTCTTTGAAACATAGATATGTAATTTCTGGATATA 900
QY 6017 AGATGCGGTTTTAACTACTTTTGCATTTTGGTATATCTTTCTTCTAATACATATGATC 6076
DB 901 AGATGCGGTTTTAACTACTTTTGCATTTTGGTATATCTTTCTTCTAATACATATGATC 960
QY 6077 AATACACTTTTGTGTTTTTAAAGAAATTTAAAACTTTATTTCAAACATCGATCAATTTTA 6136
DB 961 AATACACTTTTGTGTTTTTAAAGAAATTTAAAACTTTATTTCAAACATCGATCAATTTTA 1020
QY 6137 CTTTGTGTTCCATATTTGACTACATTTATAGGCTCACACTTTT 6178
DB 1021 CTTTGTGTTCCATATTTGACTACATTTATAGGCTCACACTTTT 1062

RESULT 9
ADL71847
ID ADL71847 standard; DNA; 1058 BP.
XX AC ADL71847;
XX DT 20-MAY-2004 (first entry)
XX DE Arabidopsis thaliana OBP3 antisense DNA #8.
XX KW Transgenic plant; Dof transcription factor; ocs binding factor;
XX KW plant size; plant stature; root growth; plant; gene; ds; OBF;
XX KW OBF binding protein; OBP3; SOB1; mouse-ear cress.
XX OS Arabidopsis thaliana.
XX PN US2004045055-A1.
XX PD 04-MAR-2004.
XX PF 28-AUG-2003; 2003US-00650249.
XX PR 28-AUG-2002; 2002US-0406657P.
```

XX	(UNIW ) UNIV WASHINGTON.
PA	Neff MM;
XX	WPI; 2004-225757/21.
DR	New transgenic plant cell, useful in producing plants with altered size
PT	and stature and with normal and healthy root growth.
XX	Claim 39; SEQ ID NO 19; 53pp; English.
XX	The invention relates to a transgenic plant transformed by a Dof
CC	transcription factor, OBP (one binding factor) binding protein (OBP3).
CC	OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic
CC	acid and polypeptides are useful in producing transgenic plants with
CC	altered size and stature and with normal and healthy root growth. The
CC	present sequence is Arabidopsis thaliana OBP3 antisense DNA.
XX	Sequence 1058 BP; 355 A; 176 C; 139 G; 388 T; 0 U; 0 Other;
QY	Query Match 14.0%; Score 1058; DB 12; Length 1058;
DB	Best Local Similarity 100.0%; Pred. No. 2.1e-154;
DB	Matches 1058; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	3045 CAACCAAGAACGATGCGTATATGATTGACTTCGAAAATAAGCAACAATACTCCTGT 3104
DB	
QY	3105 CAAATCGACACTTAATCCAAAAAGGTTAGTAATAAGTAAGAAGGCCTTTTATTATGAAA 3164
DB	
QY	3165 ACAAAAAGAAATAAAGACCTTAAGAGAAATGATGAAAAATTGAAGAGAAAAAGAGCATG 3224
DB	
QY	3225 TTATAGAAAGAAAAAGAGAGAGTAAGAGAAATTAAGAAACACAAATAATTAACAA 3284
DB	
QY	3285 AGGAAACTTCATTTCTCTCTTTATCCCAITTCAGCTCCTCCCCTCTCTCTCTCTCTCT 3344
DB	
QY	3345 CTCTCTCTCTAGATCAATCTTTCTTATAGATGATATTCACCAATATCTGGACC 3404
DB	
QY	3405 TCCTTACCCTAAAAGGATACAAGTAAGAGATTCAAAGATGGTTTTCTCATCTCTCCAGTG 3464
DB	
QY	3465 AATCAGTTTCGATTTCCAAAAATGGCAGCAGGTAAAAATTCAGTTTATGATATTTGCTAGAT 3524
DB	
QY	3525 GTTTCGATTCGTTCTCTTTTCTCCAAGCTCGATCAAGATTTTATGAAAAATTTGATGAGA 3584
DB	
QY	3585 TTTTGTTCGACAAAATTCCTAGCTATTGTGCGCGCGCATATATATTACTTATCAATATTC 3644
DB	
QY	3645 TTAGTTGATTAACCCCTTTTTTTTCTGTGCTCTTCGAAATATACGAAATATATAAAGAT 3704
DB	
QY	3705 GATTTCAAATTTGGTCTTTTTTCTCACTCAAGACTTTTTTAAAAAATATTCTTAGTTGA 3764
DB	
QY	661 GATTTCAAATTTGGTCTTTTTTCTCACTCAAGACTTTTTTAAAAAATATTCTTAGTTGA 720
DB	

QY	3765	TAAAAACCTTTTCTTGTCTTCTCCAAAGGCTTATGTATTAATGTTTTCTTTACAGGATT	3825			
Db	721	TAAAAACCTTTTCTTGTCTTCTCCAAAGGCTTATGTATTAATGTTTTCTTTACAGGATT	780			
QY	3825	AAATTTCTCTTTGGTTAGATTTTTTACACCGCCATGGAATTATCACTTCAAAAAATAAAAAA	3884			
Db	781	AAATTTCTCTTTGGTTAGATTTTTTACACCGCCATGGAATTATCACTTCAAAAAATAAAAAA	840			
QY	3885	GTTTAAAGTTACTATGACTTTTAAATCTGAGTTATTTATTCATTTTCTTTTTCGACGTTTGT	3944			
Db	841	GTTTAAAGTTACTATGACTTTTAAATCTGAGTTATTTATTCATTTTCTTTTTCGACGTTTGT	900			
QY	3945	TGAAAAACTATAATTAATCTGCAATCTTGTCAAAGTAGTCACAAATTTTTTATCTATTTC	4004			
Db	901	TGAAAAACTATAATTAATCTGCAATCTTTGTCAAGTAGTCACAAATTTTTTATCTATTTC	960			
QY	4005	TTTTGTCTCCGACCAATGTTTTCAAATCGAATCCTTTCTGTTAAAGTTGTTCTGCTTTA	4064			
Db	961	TTTTGTCTCCGACCAATGTTTTCAAATCGAATCCTTTCTGTTAAAGTTGTTCTGCTTTA	1020			
QY	4065	TTATAAACCTGAAACTAATTAGTACAAATTTATGTTAAT	4102			
Db	1021	TTATAAACCTGAAACTAATTAGTACAAATTTATGTTAAT	1058			
RESULT 10						
ADL71840						
ID	ADL71840 standard; DNA; 888 BP.					
XX	XX					
AC	ADL71840;					
XX	XX					
DT	20-MAY-2004 (first entry)					
XX	XX					
DE	Arabidopsis thaliana OBP3 antisense DNA #1.					
XX	XX					
KW	Transgenic plant; Dof transcription factor; ocs binding factor;					
KW	plant size; plant stature; root growth; plant; gene; ds; OBP;					
KW	OBP binding protein; OBP3; SOB1; mouse-ear cross.					
XX	XX					
OS	Arabidopsis thaliana.					
XX	XX					
PN	US2004045055-A1.					
XX	XX					
PD	04-MAR-2004.					
XX	XX					
PF	28-AUG-2003; 2003US-00650249.					
XX	XX					
PR	28-AUG-2002; 2002US-0406657P.					
XX	XX					
PA	(UNIW ) UNIV WASHINGTON.					
XX	XX					
FI	Neff MW;					
XX	XX					
DR	WPI; 2004-225757/21.					
XX	XX					
PT	New transgenic plant cell, useful in producing plants with altered size					
PT	and stature and with normal and healthy root growth.					
XX	XX					
PS	Claim 26; SEQ ID NO 12; 53pp; English.					
XX	XX					
CC	The invention relates to a transgenic plant transformed by a Dof					
CC	transcription factor, OBP (ocs binding factor) binding protein (OBP3).					
CC	OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic					
CC	acid and polypeptides are useful in producing transgenic plants with					
CC	altered size and stature and with normal and healthy root growth. The					
CC	present sequence is Arabidopsis thaliana OBP3 antisense DNA.					
XX	XX					
SQ	Sequence 888 BP; 263 A; 193 C; 197 G; 235 T; 0 U; 0 Other;					
Query Match						
Best Local Similarity 100.0%; Score 888; DB 12; Length 888;						
Matches 888; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						

QY 4232 ATGGTGGAAAGTGGCTCGGATCGAAAGTCCCATTCCTGGAAGCAGCTCTAAATGGCCCT 4291  
Db |||||  
1 ATGGTGGAAAGTGGCTCGGATCGAAAGTCCCATTCCTGGAAGCAGCTCTAAATGGCCCT 60  
QY 4292 AGATGTGACTCAACCAATACTAAGTTCGTCTTACTTCAATACTATAGCCTTACTCAACCT 4351  
Db |||||  
61 AGATGTGACTCAACCAATACTAAGTTCGTCTTACTTCAATACTATAGCCTTACTCAACCT 120  
QY 4352 CGCCATTTCTGCAAAACATGTCGTCTATTTGGACACGTGGCGGTTCTTTGAGGAATGTT 4411  
Db |||||  
121 CGCCATTTCTGCAAAACATGTCGTCTATTTGGACACGTGGCGGTTCTTTGAGGAATGTT 180  
QY 4412 CCGTGTGGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 4471  
Db |||||  
181 CCGTGTGGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 240  
QY 4472 GTGGTCTCGACTGATAATACTACTAGTACTTCACTCACTTCTTCGCCCCAAGTTACTCA 4531  
Db |||||  
241 GTGGTCTCGACTGATAATACTACTAGTACTTCACTCACTTCTTCGCCCCAAGTTACTCA 300  
QY 4532 AACCTTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATCCAACTTGCCTATC 4591  
Db |||||  
301 AACCTTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATCCAACTTGCCTATC 360  
QY 4592 TTGCTCTCTCCAAAGCCTTCGAGATTACAAATTCAGCAACACCTGGATTAGATTTTGGT 4651  
Db |||||  
361 TTGCTCTCTCCAAAGCCTTCGAGATTACAAATTCAGCAACACCTGGATTAGATTTTGGT 420  
QY 4652 GGAATCTCAATAAGCAACATGATAGTGGTATGAGTTCTAGTGGTGGGATCTTTGGATGCA 4711  
Db |||||  
421 GGAATCTCAATAAGCAACATGATAGTGGTATGAGTTCTAGTGGTGGGATCTTTGGATGCA 480  
QY 4712 TCGAGAATACTCCATCAACAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGA 4771  
Db |||||  
481 TCGAGAATACTCCATCAACAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGA 540  
QY 4772 TTGGTGCATCTTCAACCGGTTATATCCATTAAGTATAGAGGTAAGGAGGTTGTAATCAA 4831  
Db |||||  
541 TTGGTGCATCTTCAACCGGTTATATCCATTAAGTATAGAGGTAAGGAGGTTGTAATCAA 600  
QY 4832 GGTGATTTCTCAACAGAGTAGTATGATTTATTCATTCAGCTAATGTTTAAAGCCTTTGATG 4891  
Db |||||  
601 GGTGATTTCTCAACAGAGTAGTATGATTTATTCATTCAGCTAATGTTTAAAGCCTTTGATG 660  
QY 4892 GATTTTCTTCAGCGGGTTAGCGCCACGCAAAACAGAAATGTGAAGCGGAAGAGAT 4951  
Db |||||  
661 GATTTTCTTCAGCGGGTTAGCGCCACGCAAAACAGAAATGTGAAGCGGAAGAGAT 720  
QY 4952 GATCAGGATCGGGTAGGATGGGATGGAGTGAATTAATCTTATCAAGAACTTTTGGGT 5011  
Db |||||  
721 GATCAGGATCGGGTAGGATGGGATGGAGTGAATTAATCTTATCAAGAACTTTTGGGT 780  
QY 5012 AATATCAACATAAATCTCAGCGAGGAACAGGAATACACATCATGGGAGGTAAACAGTTCT 5071  
Db |||||  
781 AATATCAACATAAATCTCAGCGAGGAACAGGAATACACATCATGGGAGGTAAACAGTTCT 840  
QY 5072 TGGACCGGTTTCACTCCAACTCAACAGGCCATCTCTCATTTCTAA 5119  
Db |||||  
841 TGGACCGGTTTCACTCCAACTCAACAGGCCATCTCTCATTTCTAA 888

RESULT 11  
ADD30377

ID ADD30377 standard; cDNA; 795 BP.

XX ADD30377;

AC ADD30377;

XX 15-JAN-2004 (first entry)

DT 15-JAN-2004 (first entry)

DE Plant yield-related polynucleotide clone G1906.

DE ds; transcription factor; transgenic plant; growth rate; senescence;

KW seed germination rate; plant vigor; seedling vigor.

XX Arabidopsis thaliana.  
OS WO2003013227-A2.  
XX 20-FEB-2003.  
XX 09-AUG-2002; 2002WO-US025805.  
XX 09-AUG-2001; 2001US-0310847P.  
PR 19-NOV-2001; 2001US-0336049P.  
PR 11-DEC-2001; 2001US-0338692P.  
PR 14-JUN-2002; 2002US-00171468.  
XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;  
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;  
PI Broun PE;  
XX WPI: 2003-248221/24.  
DR P-PSDB; ADD30378.  
XX New plant transcription factor polynucleotides and polypeptides, useful  
PT in producing transgenic plants with commercially valuable properties,  
PI such as an alteration in a plant growth characteristic, e.g. growth rate  
PT or apomixis.  
XX Disclosure; SEQ ID NO 406; 454pp; English.  
XX The invention relates to a number of isolated Arabidopsis thaliana cDNA  
CC sequences and their encoded proteins which are especially transcription  
CC factor related cDNA s and proteins. The isolated or recombinant plant  
CC transcription factor polynucleotides and polypeptides are useful in  
CC producing transgenic plants with commercially valuable properties, i.e.  
CC modified or altered desirable traits as compared to a reference plant,  
CC such as an alteration in a plant growth characteristic, e.g. growth rate,  
CC germination rate of seeds, vigor of plants and seedlings, or leaf and  
CC flower senescence. Sequence information related to the polynucleotides  
CC and polypeptides can also be used in bioinformatic search methods. The  
CC transgenic plant is useful for growing a progeny plant from a parent  
CC plant. This sequence represents one of the cDNAs of the invention.  
XX SQ Sequence 795 BP; 236 A; 180 C; 175 G; 204 T; 0 U; 0 Other;  
SQ Query Match 9.1%; Score 692; DB 10; Length 795;  
Best Local Similarity 89.5%; Pred. No. 8.2e-98;  
Matches 795; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

QY 4232 ATGGTGGAAAGTGGCTCGGATCGAAAGTCCCATTCCTGGAAGCAGCTCTAAATGGCCCT 4291

Db 1 ATGGTGGAAAGTGGCTCGGATCGAAAGTCCCATTCCTGGAAGCAGCTCTAAATGGCCCT 60

QY 4292 AGATGTGACTCAACCAATACTAAGTTCGTCTTACTTCAATACTATAGCCTTACTCAACCT 4351

Db 61 AGATGTGACTCAACCAATACTAAGTTCGTCTTACTTCAATACTATAGCCTTACTCAACCT 120

QY 4352 CGCCATTTCTGCAAAACATGTCGTCTATTTGGACACGTGGCGGTTCTTTGAGGAATGTT 4411

Db 121 CGCCATTTCTGCAAAACATGTCGTCTATTTGGACACGTGGCGGTTCTTTGAGGAATGTT 180

QY 4412 CCGTGTGGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 4471

Db 181 CCGTGTGGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 240

QY 4472 GTGGTCTCGACTGATAATACTACTAGTACTTCACTCACTTCTTCGCCCCAAGTTACTCA 4531

Db 241 GTGGTCTCGACTGATAATACTACTAGTACTTCACTCACTTCTTCGCCCCAAGTTACTCA 300

QY 4532 AACCTTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATCCAACTTGCCTATC 4591

Db 301 AACCTTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATCCAACTTGCCTATC 360







KW plant resistance; viral infection resistance; maize plant;  
KW maize dwarf mosaic virus; MDMV; ds.

OS Synthetic.

XX Key Location/Qualifiers  
FH 119..672  
FT /\*tag= a

FT /note= "Zea mays alcohol dehydrogenase 1, intron 1  
fragment"

XX W09421796-A2.

XX 29-SEP-1994.

XX 22-MAR-1994; 94WO-US003028.

XX 24-MAR-1993; 93US-00038768.

XX (PION-) PIONEER HI-BRED INT INC.

XX (USDA ) US SEC OF AGRIC.

XX Roth BA, Townsend R, McMullen MD;

XX WPI; 1994-317016/39.

XX DNA encoding maize chlorotic dwarf virus proteins - used to provide  
XX plants with resistance to the virus and related viral infections.

XX Claim 2; Page 14-17; 40pp; English.

XX This sequence shows the nucleotide sequence for the plasmid designated  
XX pPH11406. The plasmid vector contains the gene for MCDV (maize chlorotic  
XX dwarf virus) coat protein 3 placed under control of tandem cauliflower  
XX mosaic virus 35S promoters isolated from the 1841 strain of the virus.  
XX and a polyA signal sequence obtained from the potato proteinase inhibitor  
XX II (Pin II) gene that exhibits enhancer-like activity. The chimeric gene  
XX also includes a 79 bp sequence omega' from the 5' leader region of  
XX tobacco mosaic virus that functions as a translational enhancer; and a  
XX Zea mays alcohol dehydrogenase 1, intron 1 fragment spanning nucleotides  
XX 119-672, trimmed to 557 bp with Bal 31 nuclease, which has been shown to  
XX function as an enhancer of gene expression in monocots. The DNA is used  
XX for imparting resistance to MCDV or viruses to which MCDV infection or  
XX resistance provides cross-resistance, including maize dwarf mosaic virus  
XX strain A. Any or all of the three coat protein genes from MCDV can be  
XX used to provide protection for plants. MCDV has a single, long RNA core  
XX which is shown in AAG74694. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 5033 BP; 1345 A; 1177 C; 1177 G; 1334 T; 0 U; 0 Other;

Query Match 8.2%; Score 621; DB 2; Length 5033;  
Best Local Similarity 95.1%; Pred. No. 8.5e-87;  
Matches 653; Conservative 0; Mismatches 30; Indels 4; Gaps 1;

QY 6868 AGATATCATCAATCCACTTCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGA 6927

Db 1103 AGATATCATCAATCCACTTCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGA 1044

QY 6928 TGCTCTCGTGGGTGGGTTCATCTTTGGGACCACTGTCGTAGAGGACATCTGAACGA 6987

Db 1043 TGCTCTCGTGGGTGGGTTCATCTTTGGGACCACTGTCGTAGAGGACATCTGAACGA 984

QY 6988 TAGCCTTTCTTTATCGCAATGCGCATTTCTAGAACGCACTCTCTTTCTACTGTCC 7047

Db 993 TAGCCTTTCTTTATCGCAATGCGCATTTCTAGAACGCACTCTCTTTCTACTGTCC 924

QY 7048 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATATCC 7107

Db 923 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATATCC 864

QY 7108 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTGTATTTCTTGG 7167

Db 863 TTTGTTGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTGTATTTCTTGG 804

QY 7168 AGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCT-----AGATATCATCAATCC 7223  
Db 803 AGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCTTTCTAGAGCATCAATCAATCC 744  
QY 7224 ACTTGCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGATGTTCTCTGGTGGTGG 7283  
Db 743 ACTTGCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGATGTTCTCTGGTGGTGG 684  
QY 7284 GGTCCATCTTTGGGACCACTGTCGTAGAGGACATCTTGAAGATAGCCCTTTCTTTATCG 7343  
Db 683 GGTCCATCTTTGGGACCACTGTCGTAGAGGACATCTTGAAGATAGCCCTTTCTTTATCG 624  
QY 7344 CAATGATGCGCATTTGTAGAGCCATCTCTCTTTCTACTGTCTTTCGATGAAGTGACAG 7403  
Db 623 CAATGATGCGCATTTGTAGAGCCATCTCTCTTTCTACTGTCTTTCGATGAAGTGACAG 564  
QY 7404 ATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTTACCTTTGTTGAAAAGTCTCA 7463  
Db 563 ATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTTACCTTTGTTGAAAAGTCTCA 504  
QY 7464 ATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGTGTG 7523  
Db 503 ATAGCCCTTTGCTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGTGTG 444  
QY 7524 TGCTCCACCATGTTGGGATCCACTAG 7550  
Db 443 TGCTCCACCATGTTGACGGATTTCTAG 417

#### RESULT 15

AEA40942/c

ID AEA40942 standard; DNA; 4727 BP.

XX AEA40942;

DT 28-JUL-2005 (first entry)

XX T-DNA region of pMYL28.

XX interleukin-12; transgenic plant; Nicotiana tabacum; T-DNA; pMYL28; ds.

XX Unidentified.

XX KR2004097575-A.

XX 18-NOV-2004.

XX 12-MAY-2003; 2003KR-00029910.

XX 12-MAY-2003; 2003KR-00029910.

XX (KWON/) KWON T H.

XX (YANG/) YANG M S.

XX Kwon TH, Yang MS;

XX WPI; 2005-238941/25.

XX Production method of a large amount of biologically active interleukin-12  
XX protein using plant cell suspension culture cheaply and safely.

XX Disclosure; SEQ ID NO 2; 23pp; Korean.

XX A production method of biologically active interleukin-12 protein using  
XX plant cell suspension culture is provided, thereby cheaply and safely  
XX producing a large quantity of biologically active interleukin-12 protein.  
XX The production method of biologically active interleukin-12 protein  
XX comprises the steps of: (a) providing a transgenic plant having a gene  
XX encoding p35 subunit of interleukin-12 (IL-12) and a transgenic plant  
XX having a gene encoding p40 subunit of interleukin-12 (IL-12); (b) cross-  
XX breeding the p35 containing transgenic plant with the p40 containing  
XX transgenic plant to produce a transgenic plant expressing both p35 and

CC p40; (c) isolating calluses of the transgenic plant of step (b) to  
CC produce a plant cell line(KTC 1021BP); and (d) suspension culturing the  
CC plant cell line(KTC 1021BP) to express the biologically active  
CC interleukin-12, wherein the plant is Nicotiana tabacum; gelatin may be  
CC further added into the medium in step (d) for improving production of IL-  
CC 12; and the IL-12 contains a signal peptide. The present sequence  
CC represents T-DNA region of pMYL28.  
XX  
SQ Sequence 4727 BP; 1218 A; 1135 C; 1191 G; 1183 T; 0 U; 0 Other;

Query Match	8.1%;	Score 617.2;	DB 14;	Length 4727;
Best Local Similarity	93.6%;	Pred. No. 3.3e-86;		
Matches	659;	Conservative	0;	Mismatches 33; Indels 12; Gaps 1;

  

QY	6868	AGATATCAATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCCTTTTCCACGA	6927
Db	3224	AGATATCAATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCCTTTTCCACGA	3165
QY	6928	TGTTCCCTCGTGGGTGGGTCCATCTTTGGGACCACTGTCGGTAGGCACTTTGAACGA	6987
Db	3164	TGCTCCTCGTGGGTGGGTCCATCTTTGGGACCACTGTCGGCAGAGGCACTTTGAACGA	3105
QY	6988	TAGCCCTTTCCCTTATCGCAATGATGGCAATTTGTAGAAGCACTTTCTTTTCTACTGTCC	7047
Db	3104	TAGCCCTTTCCCTTATCGCAATGATGGCAATTTGTAGGTGCCACCTTTCTTTCTACTGTCC	3045
QY	7048	TTTCGATGAAGTGA CAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCC	7107
Db	3044	TTTTCGATGAAGTGA CAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCC	2985
QY	7108	TTTGTGTGAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTTGG	7167
Db	2984	TTTGTGTGAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGG	2925
QY	7168	AGTAGACGAGAGTGTGCTGCTCCACATGTTGGGGATCTAGATATCAATCAATCCACTT	7227
Db	2924	AGTAGACGAGAGTGTGCTGCTCCACATGT-----TATCAATCAATCCACTT	2877
QY	7228	GCTTTTGAAGACGTGGTTGGAACGTCCTTTTTCACGATGTCCTCGTGGGTGGGGTC	7287
Db	2876	GCTTTTGAAGACGTGGTTGGAACGTCCTTTTTCACGATGTCCTCGTGGGTGGGGTC	2817
QY	7288	CATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCCTTCTTTATCGCAAT	7347
Db	2816	CATCTTTGGGACCACTGTCGGCAGAGGCATCTTGAACGATAGCCCTTCTTTATCGCAAT	2757
QY	7348	GATGGCAATTTGTAAGACCATCTTCTTTTCTACTGTCTTTTCGATGAAGTGACAGATAG	7407
Db	2756	GATGGCAATTTGTAAGTGCCACCTTCTTTTCTACTGTCTTTTGAAGTGACAGATAG	2697
QY	7408	CTGGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTGTTGTAAGTCTCAATAG	7467
Db	2696	CTGGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTGTTGTAAGTCTCAATAG	2637
QY	7468	CCCTCTGGTCTTCTCAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGTGCTGCT	7527
Db	2636	CCCTTTGGTCTTCTCAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGTGCTGCT	2577
QY	7528	CCACCATGTTGGGGATCCACTAGTTCTAGAGGGCGGCCACCGC	7571
Db	2576	CCACCATGTTGGCAAGCTGCTTAGGCAATACGCAACCGGCTC	2533

Search completed: December 30, 2005, 17:26:48  
Job time : 2706 secs

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).

Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.raphm** (Published\_Applications\_AA\_Main) and **.raphn** (Published\_Applications\_AA\_New).

BEST AVAILABLE COPY

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2005, 17:26:56 ; Search time 470 Seconds  
(without alignments)  
8454.579 Million cell updates/sec

Title: US-10-650-249-1

Perfect score: 7580

Sequence: 1 agctctattaattcaagaga.....gcgcaccgcgggtggagct 7580

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4172979 seqs, 262114271 residues

Total number of hits satisfying chosen parameters: 8345958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA\_New:\*  
1: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	617.2	8.1	11667	6	US-10-508-263-122
C 2	381.2	5.0	612	7	US-11-038-981A-1
C 3	379.2	5.0	1459	7	US-11-038-981A-23
C 4	377.6	5.0	3469	7	US-11-192-801-23
C 5	377.6	5.0	3754	7	US-11-192-801-15
C 6	376.2	5.0	1467	7	US-11-038-981A-21
C 7	374.6	4.9	1467	7	US-11-038-981A-22
C 8	373.6	4.9	524	7	US-11-038-981A-6
C 9	367.2	4.8	4149	7	US-11-192-801-13
C 10	303.6	4.0	835	6	US-10-391-414-10
C 11	118	1.6	1935	7	US-11-038-981A-26
C 12	116.2	1.5	332	7	US-11-038-981A-5
C 13	116.2	1.5	1273	7	US-11-038-981A-18
C 14	116.2	1.5	1281	7	US-11-038-981A-20
C 15	116.2	1.5	1935	7	US-11-038-981A-19
C 16	116.2	1.5	1935	7	US-11-038-981A-25
C 17	116.2	1.5	1939	7	US-11-038-981A-24
C 18	116.2	1.5	1963	7	US-11-038-981A-27
C 19	116.2	1.5	1963	7	US-11-038-981A-28
C 20	116.2	1.5	1971	7	US-11-038-981A-29
C 21	101.6	1.3	986	6	US-10-927-641-40
C 22	88.6	1.2	762	6	US-10-509-691-1
C 23	78.2	1.0	173602	7	US-11-121-086-25

C 24	69.2	0.9	171486	7	US-11-121-086-105
C 25	68	0.9	171486	7	US-11-121-086-105
C 26	66.6	0.9	173602	7	US-11-121-086-25
C 27	65.8	0.9	5666	6	US-10-240-708-29
C 28	65.2	0.9	5152	6	US-10-240-708-47
C 29	63	0.8	49979	6	US-10-995-561-13443
C 30	62.6	0.8	5562	6	US-10-240-708-63
C 31	62.6	0.8	105550	6	US-10-995-561-13235
C 32	62	0.8	8093	6	US-10-240-708-31
C 33	61.6	0.8	1073	7	US-11-038-981A-16
C 34	61.4	0.8	119036	6	US-10-995-561-13314
C 35	61.2	0.8	6317	6	US-10-240-708-11
C 36	60.8	0.8	20317	6	US-10-995-561-13460
C 37	60.6	0.8	8306	6	US-10-240-708-50
C 38	60.2	0.8	20317	6	US-10-995-561-13460
C 39	59.6	0.8	5501	6	US-10-240-708-38
C 40	59.6	0.8	6866	6	US-10-240-708-20
C 41	59.4	0.8	240	7	US-11-038-981A-2
C 42	59.4	0.8	1176	7	US-11-038-981A-11
C 43	59.4	0.8	1184	7	US-11-038-981A-9
C 44	59.4	0.8	1184	7	US-11-038-981A-10
C 45	59	0.8	125	7	US-11-038-981A-4

ALIGNMENTS

RESULT 1  
US-10-508-263-122/c  
; Sequence 122, Application US/10508263  
; GENERAL INFORMATION:  
; APPLICANT: BASF Plant Science GmbH  
; TITLE OF INVENTION: Constructs and methods for regulating gene expression  
; FILE REFERENCE: 53262-20085.00  
; CURRENT APPLICATION NUMBER: US/10/508,263  
; CURRENT FILING DATE: 2004-09-20  
; NUMBER OF SEQ ID NOS: 126  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 122  
; LENGTH: 11667  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Description of the artificial sequence: suppression  
; OTHER INFORMATION: construct 2 p3300.1-Toc159-GFP-RNAI  
US-10-508-263-122

Query Match	8.1%	Score 617.2;	DB 6;	Length 11667;
Best Local Similarity	93.6%	Pred. No. 1.4e-96;		
Matches 659;	Conservative 0;	Mismatches 33;	Indels 12;	Gaps 1;
QY	6868	AGATATACATCAATCCACCTTGTGTTGAAGACGTGGTTGGAAACGTCTTCTTTTTCACCA	6927	
DB	1112	AGATATACATCAATCCACCTTGTGTTGAAGACGTGGTTGGAAACGTCTTCTTTTTCACCA	1053	
QY	6928	TGTTCTCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTGAACCA	6987	
DB	1052	TGCTCTCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTGAACCA	993	
QY	6988	TAGCCTTTCTTTTATCGCAATGATGGCATTTGTAGAACCATCTTCTTTTCTACTGTCC	7047	
DB	992	TAGCCTTTCTTTTATCGCAATGATGGCATTTGTAGGTGCCACCTTCTTTTCTACTGTCC	933	
QY	7048	TTTCGATCAAGTCAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC	7107	
DB	932	TTTGTATGAAGTCAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC	873	
QY	7108	TTTGTGTAAAAGTCTCAATAGCCCTCTGCTCTTCTTGAGACTGTATCTTTGATATTCTGG	7167	
DB	872	TTTGTGTAAAAGTCTCAATAGCCCTTGGTCTTCTTGAGACTGTATCTTTGATATTCTGG	813	
QY	7168	AGTAGACGAGGTGTGCTGCCACCATGTTGGGATCTAGATATCATCATCAATCACTT	7227	

Db 812 AGTAGACGAGAGTGTCTGCTCCACCATG-----TATCATCATCAATCCACTT 765  
Qy 7228 GCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTC 7287  
Db 764 GCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTC 705  
Qy 7288 CATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCCTTTCCTTTATCGCAAT 7347  
Db 704 CATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCCTTTCCTTTATCGCAAT 645  
Qy 7348 GATGSCATTTGTAGAAGCCATCTCTCTTTCTACTGTCTCTTTCGATGAAGTGACAGATAG 7407  
Db 644 GATGSCATTTGTAGTGTCCACCTTCTCTTTCTACTGTCTCTTTCGATGAAGTGACAGATAG 585  
Qy 7408 CTGGSCAATGGAATCCGAGGAGTTTCCGATATPACCTTTGTGAAAAGTCTCAATAG 7467  
Db 584 CTGGSCAATGGAATCCGAGGAGTTTCCGATATPACCTTTGTGAAAAGTCTCAATAG 525  
Qy 7468 CCTCTGTCTTCTGAGACTGTATCTTTTGATATTTCTTGAGTAGACGAGTGTCTGCT 7527  
Db 524 CCTTTGTCTTCTGAGACTGTATCTTTTGATATTTCTTGAGTAGACGAGTGTCTGCT 465  
Qy 7528 CCACCATGTTGGGATCCACTAGTTCTAGAGCGGCCGCCACGC 7571  
Db 464 CCACCATGTTGGCAAGCTGCTCTAGCCCAATAGCCAAACCGCTC 421

## RESULT 2

US-11-038-981A-1/c

; Sequence 1, Application US/11038981A

; Publication No. US20050283856A1

; GENERAL INFORMATION:

; APPLICANT: Conner, Timothy W.

; APPLICANT: Flasiński, Stanisław

; APPLICANT: Pang, Sheng Z

; APPLICANT: You, Jinsong

; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS

; FILE REFERENCE: 38-21(51446)B

; CURRENT APPLICATION NUMBER: US/11/038,981A

; CURRENT FILING DATE: 2005-01-20

; PRIOR APPLICATION NUMBER: 60/537,793

; PRIOR FILING DATE: 2004-01-20

; NUMBER OF SEQ ID NOS: 35

; SEQ ID NO 1

; LENGTH: 612

; TYPE: DNA

; ORGANISM: Cauliflower mosaic virus

US-11-038-981A-1

Query Match 5.0%; Score 381.2; DB 7; Length 612;  
Best Local Similarity 81.5%; Pred. No. 1.2e-56;  
Matches 489; Conservative 0; Mismatches 33; Indels 78; Gaps 1;

Qy 6190 AGATATCATCAATCCACTGCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGA 6249  
Db 526 AGATATCATCAATCCACTGCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGA 467  
Qy 6250 TGTCTCTGTGGGTGGGGTCCATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGA 6309  
Db 466 TGTCTCTGTGGGTGGGGTCCATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGA 407  
Qy 6310 TAGCCCTTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTTCTACTGTCC 6369  
Db 406 TGGCCCTTCTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTTCTACTATCT 347  
Qy 6370 TTTGATGAAGTGACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429  
Db 346 TCACAATAAAGTGACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 287  
Qy 6430 TTTGTTGAAAAGTCTCAATAGCCCTCTGTCTTCTGAGACTGTATCTTTTGATATTCTTGG 6489  
Db 286 TTTGTTGAAAAGTCTCAATCGGAC-----CATCGACCA 263

Qy 6490 AGTAGACGAGAGTGTCTGCTCCACCATGTTGGGGATCTAGATATCATCAATCCACTT 6549  
Db 262 -----CATCATCAATCCACTT 245  
Qy 6550 GCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTC 6609  
Db 244 GCTTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTC 185  
Qy 6610 CATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCCTTTCCTTTATCGCAAT 6669  
Db 184 CATCTTTGGGACCACTGTCTGGTAGAGGCATCTTCAACGATGCGCTTTCCTTTATCGCAAT 125  
Qy 6670 GATGSCATTTGTAGAAGCCATCTCTCTTTCTACTGTCTCTTTCGATGAAGTGACAGATAG 6729  
Db 124 GATGSCATTTGTAGAGCCACTTCTCTTTTCCACTATCTTCACAATAAAGTGACAGATAG 65  
Qy 6730 CTGGGCAATGGAATCCGAGGAGTTTCCGATATPACCTTTGTGAAAAGTCTCAATAG 6789  
Db 64 CTGGGCAATGGAATCCGAGGAGTTTCCGATATPACCTTTGTGAAAAGTCTCAATCG 5

## RESULT 3

US-11-038-981A-23/c

; Sequence 23, Application US/11038981A

; Publication No. US20050283856A1

; GENERAL INFORMATION:

; APPLICANT: Conner, Timothy W.

; APPLICANT: Flasiński, Stanisław

; APPLICANT: Pang, Sheng Z

; APPLICANT: You, Jinsong

; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS

; FILE REFERENCE: 38-21(51446)B

; CURRENT APPLICATION NUMBER: US/11/038,981A

; CURRENT FILING DATE: 2005-01-20

; PRIOR APPLICATION NUMBER: 60/537,793

; PRIOR FILING DATE: 2004-01-20

; NUMBER OF SEQ ID NOS: 35

; SEQ ID NO 23

; LENGTH: 1459

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Artificial promoter sequence

US-11-038-981A-23

Query Match 5.0%; Score 379.2; DB 7; Length 1459;  
Best Local Similarity 80.4%; Pred. No. 3.6e-56;  
Matches 492; Conservative 0; Mismatches 43; Indels 77; Gaps 1;

Qy 6174 CTTTGTCTTCGATCTAGATATCATCAATCCACTGCTTTGAAGACGTGGTTGGAACG 6233  
Db 1319 CTTTGAAGACGAGGAGGATCATCAATCCACTTCTTTGAAGACGTGGTTGGAACG 1260  
Qy 6234 TCTTCTTTTCCACGATGTTCTCTGTGGTGGGGTCCATCTTTGGGACCACTGTCTGCTA 6293  
Db 1259 TCTTCTTTTCCACGATGTTCTCTGTGGTGGGGTCCATCTTTGGGACCACTGTCTGCTA 1200  
Qy 6294 GAGGCATCTTGAACGATAGCCCTTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCT 6353  
Db 1199 GAGGCATCTTCAACGATGGCTTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCT 1140  
Qy 6354 TCTTTTCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGCAATGGAATCCGAGGAGG 6413  
Db 1139 TCTTTTCTCACTATCTTCACAATAAAGTGACAGATAGCTGGCAATGGAATCCGAGGAGG 1080  
Qy 6414 TTTCCGATATTACCTTTTGTGAAAAGTCTCAATAGCCCTCTGTCTTCTCTGAGACTGTA 6473  
Db 1079 TTTCCGATATTACCTTTTGTGAAAAGTCTCA----- 1047  
Qy 6474 TCTTTGATATTCTTGGAGTAGACGAGAGTGTCTCCACCATGTTGGGAGTCTAGATA 6533  
Db 1046 -----CATCGACCA 1037



QY 6534 TCACATCAATCCACTTGTCTTGAAGACGTGGTTGGAACGTCTCTTTCTTTTCCACGATGTTTC 6593  
|||||  
Db 1036 TCACATCAATCCACTTGTCTTGAAGACGTGGTTGGAACGTCTCTTTCTTTTCCACGATGTTTC 977  
|||||  
QY 6594 CTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCC 6653  
976 CTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCC 917  
Db 6654 TTTCTCTTATCGCAATGATGGCAATTTGGAAGCCATCTTCTCTTCTACTGTCTCTTTTCG 6713  
|||||  
Db 916 TTTCTCTTATCGCAATGATGGCAATTTGGAAGCCATCTTCTCTTCTACTGTCTCTTTTCACA 857  
QY 6714 ATGAAGTAGACAGATAGCTGGCAATCGGAATCGAGAGGTTTCCCGATATTTACCTTTGT 6773  
856 ATAAAGTAGACAGATAGCTGGCAATCGGAATCGAGAGGTTTCCCGATATTTACCTTTGT 797  
QY 6774 TGAAAAGTCTCA 6785  
796 TGAAAAGTCTCA 785  
Db

## RESULT 4

US-11-192-801-23/c  
; Sequence 23, Application US/11192801  
; Publication No. US20050273882A1  
; GENERAL INFORMATION:  
; APPLICANT: Romano, Charles P.  
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants  
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn  
; CURRENT APPLICATION NUMBER: US/11/192,801  
; PRIOR FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US/10/232,665  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: US/09/377,466  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 3469  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: expression  
; OTHER INFORMATION: cassette  
; NAME/KEY: promoter  
; LOCATION: (25)..(640)  
; FEATURE:  
; OTHER INFORMATION: P-CaMV.35S  
; NAME/KEY: 5'UTR  
; LOCATION: (664)..(734)  
; FEATURE:  
; OTHER INFORMATION: L-Ta.hcb1  
; NAME/KEY: intron  
; LOCATION: (748)..(1238)  
; OTHER INFORMATION: I-Os.Act1  
; NAME/KEY: CDS  
; LOCATION: (1241)..(3199)  
; OTHER INFORMATION: Cry3Bb1 variant 11231mv2  
; NAME/KEY: terminator  
; LOCATION: (3217)..(3450)  
; OTHER INFORMATION: T-Ta.hsp17  
US-11-192-801-23

Query Match 5.0%; Score 377.6; DB 7; Length 3469;  
Best Local Similarity 81.4%; Pred. No. 8.7e-56;  
Matches 485; Conservative 34; Indels 77; Gaps 1;  
QY 6190 AGATATCATCAATCCACTTGTCTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGA 6249  
|||||

Db 553 AGATATCATCAATCCACTTGTCTTGAAGACGTGGTTGGAACGTCTCTTTTCCACGA 494  
QY 6250 TGTCTCTCGTGGTGGGGTCCATCTTTGGGACCACTGTGTGGTAGAGGCATCTTGAACGA 6309  
493 TGTCTCTCGTGGTGGGGTCCATCTTTGGGACCACTGTGTGGTAGAGGCATCTTGAACGA 434  
QY 6310 TAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCC 6369  
433 TGGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCCACTATCT 374  
QY 6370 TTTCCATGAAGTAGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429  
373 TCACAAATAAGTAGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCC 314  
QY 6430 TTTGTGTAAAAGTCTCAATAGCCCTCTCGTCTTCTTGAGACTGTATCTTTGTATATTCTTGG 6489  
313 TTTGTGTAAAAGTCTCA ----- 297  
QY 6490 AGTAGACGAGAGTGTCTGTCTCCACCATGTTGGGATCTAGATATCATCAATCAATCCACTT 6549  
296 -----CATCGGACCATCAATCAATCCACTT 271  
QY 6550 GCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGATGTTCTCTGGTGGGGTCC 6609  
270 GCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGATGTTCTCTGGTGGGGTCC 211  
QY 6610 CATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCCTTTCTTTATCGCAAT 6669  
210 CATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCCTTTCTTTATCGCAAT 151  
QY 6670 GATGGCATTTGTAGAGGCCATCTTCTTCTACTGTCTTTCCGATAGTGAAGTACAGATAG 6729  
150 GATGGCATTTGTAGAGGCCATCTTCTTCTTCCACTATCTTCAACAATAAGTACAGATAG 91  
QY 6730 CTGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTTGTGAAAAGTCTCA 6785  
90 CTGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTTGTGAAAAGTCTCA 35  
Db

## RESULT 5

US-11-192-801-15/c  
; Sequence 15, Application US/11192801  
; Publication No. US20050273882A1  
; GENERAL INFORMATION:  
; APPLICANT: Romano, Charles P.  
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants  
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn  
; CURRENT APPLICATION NUMBER: US/11/192,801  
; PRIOR FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US/10/232,665  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: US/09/377,466  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 3754  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: expression  
; OTHER INFORMATION: cassette  
; NAME/KEY: promoter  
; LOCATION: (25)..(640)  
; OTHER INFORMATION: P-CaMV.35S  
; NAME/KEY: intron  
; LOCATION: (669)..(1472)  
; OTHER INFORMATION: I-Zm.Hsp70  
; NAME/KEY: CDS  
; LOCATION: (1490)..(3448)  
;



FEATURE:  
; OTHER INFORMATION: Artificial promoter sequence  
US-11-038-981A-22

Query Match 4.9%; Score 374.6; DB 7; Length 1467;  
Best Local Similarity 81.3%; Pred. No. 2.2e-55;  
Matches 482; Conservative 0; Mismatches 34; Indels 77; Gaps 1;  
  
QY 6193 TATCATCAATCCACTTCTTGAAGACGTGGTTGGAACGTCCTTTTCCACGATGT 6252  
DB 926 TATCATCAATCCACTTCTTGAAGACGTGGTTGGAACGTCCTTTTCCACGATGC 867  
  
QY 6253 TCCTCGTGGGTGGGGTCCATCTTTGGACCACTGTCGGTAGAGCACTTTGGAAGATAG 6312  
DB 866 TCCTCGTGGGTGGGGTCCATCTTTGGACCACTGTCGGTAGAGCACTTTCAACGATGG 807  
  
QY 6313 CTTTCCCTTTATCGCAATGATGGCACTTTGTAGAAGCCATCTTCCCTTTTCTACTGTCCTTT 6372  
DB 806 CTTTCCCTTTATCGCAATGATGGCACTTTGTAGAAGCCATCTTCCCTTTTCTACTGTCCTTT 747  
  
QY 6373 CGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATATACCCCTTT 6432  
DB 746 CAATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATATACCCCTTT 687  
  
QY 6433 GTTGAAGAGTCTCAATAGACCTCTCGTCTTCTGAGACTGTATCTTTGATATATCTTGGAGT 6492  
DB 686 GTTGAAGAGTCTCA----- 673  
  
QY 6493 AGACGAGAGTGTGCGTCTCCACCATGTTGGGATCTAGATATCATCAATCCACTTCTCT 6552  
DB 672 -----CATCGGACCATCATCAATCCACTTCTCT 644  
  
QY 6553 TTGAAGACGTGGTTGGAACGTCCTTTTCCACGATGTTCTCGTGGGTGGGGTCCAT 6612  
DB 643 TTGAAGACGTGGTTGGAACGTCCTTTTCCACGATGTTCTCGTGGGTGGGGTCCAT 584  
  
QY 6613 CTTTGGGACCACTGTCGGTAGAGGCACTTTGGAACGATAGCTTCTTTATTCGCAATGAT 6672  
DB 583 CTTTGGGACCACTGTCGGTAGAGGCACTTTGGAACGATAGCTTCTTTATTCGCAATGAT 524  
  
QY 6673 GGCATTTCTAGAAGCCATCTTCTTCTACTGTCCTTTGATGAAGTGACAGATAGCTG 6732  
DB 523 GGCATTTCTAGAAGCCATCTTCTTCTACTGTCCTTTGATGAAGTGACAGATAGCTG 464  
  
QY 6733 GGCATGAATCCGAGGAGTTTCCGATATATACCCCTTTGTTGAAAAGTCTCA 6785  
DB 463 GGCATGAATCCGAGGAGTTTCCGATATATACCCCTTTGTTGAAAAGTCTCA 411

RESULT 8  
US-11-038-981A-6/c  
; Sequence 6, Application US/11038981A  
; Publication No. US20050283856A1  
; GENERAL INFORMATION:  
; APPLICANT: Conner, Timothy W.  
; APPLICANT: Flasiński, Stanisław  
; APPLICANT: Pang, Sheng Z  
; APPLICANT: You, Jinsong  
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS  
; FILE REFERENCE: 38-21(51446)B  
; CURRENT APPLICATION NUMBER: US/11/038,981A  
; CURRENT FILING DATE: 2005-01-20  
; PRIOR APPLICATION NUMBER: 60/537,793  
; PRIOR FILING DATE: 2004-01-20  
; NUMBER OF SEQ ID NOS: 35  
; SEQ ID NO 6  
; LENGTH: 524  
; TYPE: DNA  
; ORGANISM: Cauliflower mosaic virus  
US-11-038-981A-6

Query Match 4.9%; Score 373.6; DB 7; Length 524;  
Best Local Similarity 81.2%; Pred. No. 2.4e-55;

Matches 481; Conservative 0; Mismatches 34; Indels 77; Gaps 1;  
  
QY 6194 ATCATCATCAATCCACTTCTTGAAGACGTGGTTGGAACGTCCTTTTCCACGATGTT 6253  
DB 524 ATCATCATCAATCCACTTCTTGAAGACGTGGTTGGAACGTCCTTTTCCACGATGCT 465  
  
QY 6254 CCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCACTTTGAAACGATAGC 6313  
DB 464 CCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCACTTTCAACGATGGC 405  
  
QY 6314 CTTTCCCTTTATCGCAATGATGGCACTTTGTAGAAGCCATCTTCTCTTTTCTACTGTCCTTTTC 6373  
DB 404 CTTTCCCTTTATCGCAATGATGGCACTTTGTAGAAGCCATCTTCTCTTTTCTACTGTCCTTTTC 345  
  
QY 6374 GATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATATACCCCTTTG 6433  
DB 344 AATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATATACCCCTTTG 285  
  
QY 6434 TTGAAAAGTCTCAATAGACCTCTCGTCTTCTGAGACTGTATCTTTGATATATCTTGGAGTA 6493  
DB 284 TTGAAAAGTCTCA----- 272  
  
QY 6494 GACGAGAGTGTGCTGCTCCACCATGTTGGGATCTAGATATCATCAATCCACTTCTGCTT 6553  
DB 271 -----CATCGGACCATCATCAATCCACTTCTGCTT 242  
  
QY 6554 TGAAGACGTGGTTGGAACGTCCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATC 6613  
DB 241 TGAAGACGTGGTTGGAACGTCCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATC 182  
  
QY 6614 TTTGGGACCACTGTCGGTAGAGGCACTTTGGAACGATAGCTTCTCTTTATTCGCAATGATG 6673  
DB 181 TTTGGGACCACTGTCGGTAGAGGCACTTTGGAACGATAGCTTCTCTTTATTCGCAATGATG 122  
  
QY 6674 GCATTTGTAGAAGCCATCTTCTTCTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGG 6733  
DB 121 GCATTTGTAGAAGCCATCTTCTTCTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGG 62  
  
QY 6734 GCAATGGAATCCGAGGAGTTTCCGATATATACCCCTTTGTTGAAAAGTCTCA 6785  
DB 61 GCAATGGAATCCGAGGAGTTTCCGATATATACCCCTTTGTTGAAAAGTCTCA 10

RESULT 9  
US-11-192-801-13/c  
; Sequence 13, Application US/11192801  
; Publication No. US20050273882A1  
; GENERAL INFORMATION:  
; APPLICANT: Romano, Charles P.  
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants  
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn  
; CURRENT APPLICATION NUMBER: US/11/192,801  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US/10/232,665  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: US/09/377,466  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 4149  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: expression  
; OTHER INFORMATION: cassette  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (25)..(640)  
; OTHER INFORMATION: P-CaMV.35S  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (669)..(1472)

```
; OTHER INFORMATION: I-Zm.Hsp70
; FEATURE:
; NAME/KEY: transit_peptide
; LOCATION: (1489)..(1635)
; OTHER INFORMATION: amino terminal TS-Zm.rbcS
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1636)..(1798)
; OTHER INFORMATION: I-Zm.rbcS
; FEATURE:
; NAME/KEY: transit_peptide
; LOCATION: (1799)..(1885)
; OTHER INFORMATION: carboxy terminus TS-Zm.rbcS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1885)..(3843)
; OTHER INFORMATION: Cry3Bb1 variant v11231
; FEATURE:
; NAME/KEY: terminator
; LOCATION: (3871)..(4127)
; OTHER INFORMATION: T-AGRTu.nos 3' transcription termination and
; OTHER INFORMATION: polyadenylation sequence
US-11-192-801-13

Query Match 4.8%; Score 367.2; DB 7; Length 4149;
Best Local Similarity 81.4%; Pred. No. 5.5e-54;
Matches 485; Conservative 0; Mismatches 33; Indels 78; Gaps 2;

QY 6190 AGATATCACATCAATCCACTTGTCTTTGAAGACGTTGGTGGAAACGCTCTCTTTTCCACGA 6249
DB 552 AGATATCACATCAATCCACTTGTCTTTGAAGACGTTGGTGGAAACGCTCTCTTTTCCACGA 493
QY 6250 TGTTCCTCGTGGTGGGGTGCATCTTTGGGACCACTGTGCGGAGAGCATCTTGAACGA 6309
DB 492 TGTCTCTCGTGGTGGGGTGCATCTTTGGGACCACTGTGCGGAGAGCATCTTGAACGA 433
QY 6310 TAGCCTTTCCCTTATCGCAATGGCATTTGTAGAAGCCATCTTCCCTTTCTACTGTCC 6369
DB 432 TGGCCTTTCCCTTATCGCAATGGCATTTGTAGAAGCCATCTTCCCTTTCTACTGTCT 374
QY 6370 TTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCC 6429
DB 373 TCACATTAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCC 314
QY 6430 TTGTGTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATCTTTGG 6489
DB 313 TTGTGTGAAAGTCTCA----- 297
QY 6490 AGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGATCTAGATATCACATCAATCCACTT 6549
DB 296 -----CATCGGACCATCACATCAATCCACTT 271
QY 6550 GCTTTTGAAGACGTGTTGGAAGCTCTCTTTTCCACGATGTTCTCGTGGTGGGGTGC 6609
DB 270 GCTTTTGAAGACGTGTTGGAAGCTCTCTTTTCCACGATGTTCTCGTGGTGGGGTGC 211
QY 6610 CATCTTTGGGACCACTGTGGTAGAGGCATCTTGAACGATAGCCCTTCTTTATCGCAAT 6669
DB 210 CATCTTTGGGACCACTGTGGTAGAGGCATCTTGAACGATAGCCCTTCTTTATCGCAAT 151
QY 6670 GATGGCATTTGTAAGGCCATCTTCTTTTCTACTGTCTTTTCGATGAAGTGACAGATAG 6729
DB 150 GATGGCATTTGTAAGGCCATCTTCTTTTCTACTGTCTTTTCGATGAAGTGACAGATAG 91
QY 6730 CTGGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTGTTGAAAGTCTCA 6785
DB 90 CTGGGCAATGGAATCCGAGAGGTTTCCGATATTACCTTTGTTGAAAGTCTCA 35

RESULT 10
US-11-391-414-10/c
; Sequence 10, Application US/10391414
; Publication No. US20050278799A1

; GENERAL INFORMATION:
; APPLICANT: SHINOZAKI, Kazuko
; APPLICANT: KATSURA, Koji
; APPLICANT: ITO, Yusuke
; TITLE OF INVENTION: Stress Induced Promoter Derived From Rice
; FILE REFERENCE: 382.1041
; CURRENT APPLICATION NUMBER: US/10/391.414
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: JP 2002-377316
; PRIOR FILING DATE: 2002-12-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 835
; TYPE: DNA
; ORGANISM: Cauliflower mosaic virus
US-10-391-414-10

Query Match 4.0%; Score 303.6; DB 6; Length 835;
Best Local Similarity 91.7%; Pred. No. 2.5e-43;
Matches 321; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 7207 AGATATCACATCAATCCACTTGTCTTTGAAGACGTTGGTGGAAACGCTCTTTTCCACGA 7266
DB 744 AGATATCACATCAATCCACTTGTCTTTGAAGACGTTGGTGGAAACGCTCTTTTCCACGA 685
QY 7267 TGTTCCTCGTGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTTGAACGA 7326
DB 684 TGTCTCTCGTGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTTGAACGA 625
QY 7327 TAGCCTTTCCCTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTCTACTGTCC 7386
DB 624 TGGCCTTTCCCTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTCTACTGTCT 565
QY 7387 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCC 7446
DB 564 TCACATTAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCC 505
QY 7447 TTTGTGAAAGTCTCAATAGCCCTCTGCTTCTGAGACTGTATCTTTGATATCTTTGG 7506
DB 504 TTTGTGAAAGTCTCAATAGCCCTCTGCTTCTGAGACTGTATCTTTGATATCTTTGG 445
QY 7507 AGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGATCCCACTAGTCTTAG 7556
DB 444 AGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGATCCCACTAGTCTTAG 395

RESULT 11
US-11-038-981A-26
; Sequence 26, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanislaw
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038.981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 26
; LENGTH: 1935
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-26

Query Match 1.6%; Score 118; DB 7; Length 1935;
Best Local Similarity 78.0%; Pred. No. 1.8e-11;
```

```
Matches 142; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 6468 ACTGATCTTTGATATCTTGGAGTAGACGAGGAGTGTGCTCCACCAATGTTGGGATC 6527
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
497 ACAGTAATATTTGTAAGCTTCGAAGGATAGTGGGATTTGCGTCATCCCTTACGTCAGTG 556
QY 6528 TAGATATCATCATCAATCCACTTTCCTTTGAAGAGCGTGTGGAACTCTCTTTTCCACG 5587
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
557 GAGATATAACATCAATCCACTTTCCTTTGAAGAGCGTGTGGAACTCTCTTTTCCACG 616
QY 6588 ATGTTCTCTGCTGGGTTGGGTTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACG 6647
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
617 ATGCTCTCTGCTGGGTTGGGTTCCATCTTTGGGACCACTGTCGGCAGAGGCATCTTCAACG 676
QY 6648 AT 6649
Db |||||
677 AT 678

RESULT 12
US-11-038-981A-5/c
; Sequence 5, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038, 981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537, 793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 5
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Cauliflower mosaic virus
US-11-038-981A-5

Query Match 1.5%; Score 116.2; DB 7; Length 332;
Best Local Similarity 97.5%; Pred. No. 2.1e-11;
Matches 118; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 6190 AGATATCATCAATCAATCCACTTCTTTGAAGAGCGTGTGGAACTCTCTTTTCCACGA 6249
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
122 AGATATCATCAATCAATCCACTTCTTTGAAGAGCGTGTGGAACTCTCTTTTCCACGA 63
QY 6250 TGTTCCTCTGCTGGGTTGGGTTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
Db 62 TGTCTCTCTGCTGGGTTGGGTTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 3
QY 6310 T 6310
Db 2 T 2

RESULT 13
US-11-038-981A-18/c
; Sequence 18, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038, 981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537, 793
; PRIOR FILING DATE: 2004-01-20
```

```
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 18
; LENGTH: 1273
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-18

Query Match 1.5%; Score 116.2; DB 7; Length 1273;
Best Local Similarity 97.5%; Pred. No. 3.1e-11;
Matches 118; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 6190 AGATATCATCAATCAATCCACTTCTTTGAAGAGCGTGTGGAACTCTCTTTTCCACGA 6249
Db 134 AGATATCATCAATCAATCCACTTCTTTGAAGAGCGTGTGGAACTCTCTTTTCCACGA 75
QY 6250 TGTTCCTCTGCTGGGTTGGGTTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
Db 74 TGTCTCTCTGCTGGGTTGGGTTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 15
QY 6310 T 6310
Db 14 T 14

RESULT 14
US-11-038-981A-20/c
; Sequence 20, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038, 981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537, 793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 20
; LENGTH: 1273
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-20

Query Match 1.5%; Score 116.2; DB 7; Length 1273;
Best Local Similarity 97.5%; Pred. No. 3.1e-11;
Matches 118; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 6190 AGATATCATCAATCAATCCACTTCTTTGAAGAGCGTGTGGAACTCTCTTTTCCACGA 6249
Db 898 AGATATCATCAATCAATCCACTTCTTTGAAGAGCGTGTGGAACTCTCTTTTCCACGA 839
QY 6250 TGTTCCTCTGCTGGGTTGGGTTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
Db 838 TGTCTCTCTGCTGGGTTGGGTTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 779
QY 6310 T 6310
Db 778 T 778

RESULT 15
US-11-038-981A-19/c
; Sequence 19, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
```

Tue Jan 3 11:29:08 2006

```
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 19
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-19

Query Match          1.5%; Score 116.2; DB 7; Length 1281;
Best Local Similarity 97.5%; Pred No. 3.2e-11;
Matches 118; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGGCTTTGAAGACGTGGTTGGAACTCTTTTCCACGA 6249
Db 524 AGATATCACATCAATCCACTTGGCTTTGAAGACGTGGTTGGAACTCTTTTCCACGA 465

QY 6250 TGTTCCTCGTGGGTGGGGTCCATCTTTGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
Db 464 TGCTCCTCGTGGGTGGGGTCCATCTTTGGGACCACCTGTCGGCAGAGGCATCTTCAACGA 405

QY 6310 T 6310
Db 404 T 404
```

Search completed: December 31, 2005, 06:31:43  
Job time : 477 secs









Db	4681	TATGAGTCTTAGTGGTGGGATCTTGGATGCGATGGAGATACCTCCATCACAAACAGCTCA	4740	5761	GAGAAATTATGCACCTAAATTCAGACTAATCCCCCAAAATTTTCAGAAAATTTATGTAATTTTTG	5820
Qy	4741	GCAATTCCTTTCTTTGATCAACATACACCGGATGGTGCAATCTTCAAAACGGTTATATCC	4800	5821	CGATTTAATAATTGTTTCACAATCATATAGCCAACTAACTAAATTTGAAAACAGCAATGGAA	5880
Db	4741	GCAATTCCTTTCTTTGATCAACATACACCGGATGGTGCAATCTTCAAAACGGTTATATCC	4800	5821	CGATTTAATAATTGTTTCACAATCATATAGCCAACTAACTAAATTTGAAAACAGCAATGGAA	5880
Qy	4801	ATTACTAGAAGTAAGGAGGTGTTAAATCAAGGTGATTTCTCAACAGAAGAGTAGTGATTA	4860	5881	TGACTGAAACCATGCATAATCTCTCAAGTCTCAACCTATGAAGAATCATGTAAACCAATPAG	5940
Db	4801	ATTACTAGAAGTAAGGAGGTGTTAAATCAAGGTGATTTCTCAACAGAAGAGTAGTGATTA	4860	5881	TGACTGAAACCATGCATAATCTCTCAAGTCTCAACCTATGAAGAATCATGTAAACCAATPAG	5940
Qy	4861	TTCCCAATCAGCTAAATGTTTAAGCCCTCGATGATTTCTTCAGCGCGGGTTAGCGCCAC	4920	5941	ACTATCATCATGATTAGTAAATGATATTAATGATATCTTTTGAACAATAGATATGTC	6000
Db	4861	TTCCCAATCAGCTAAATGTTTAAGCCCTCGATGATTTCTTCAGCGCGGGTTAGCGCCAC	4920	5941	ACTATCATCATGATTAGTAAATGATATTAATGATATCTTTTGAACAATAGATATGTC	6000
Qy	4921	GCAAAACAAGAAATGTGAAGCGGAGAGAAATGATCAGGATCGGGTAGGGATGG	4980	6001	ATTTTATCTGGATATAAAGATGGCGTTTAACTACTTTTGCAATTTTGTGTTATATCTTTCT	6060
Db	4921	GCAAAACAAGAAATGTGAAGCGGAGAGAAATGATCAGGATCGGGTAGGGATGG	4980	6001	ATTTTATCTGGATATAAAGATGGCGTTTAACTACTTTTGCAATTTTGTGTTATATCTTTCT	6060
Qy	4981	AGTGAATAACTTATCAAGAAACTTTTGGGTAAATATCAACATAAATCTCAGGCAGGAACGA	5040	6061	TCTAATACATATGATCAATACACTTTTGTGTTTAAAAAGAAATTAATACTTATTTCAAAC	6120
Db	4981	AGTGAATAACTTATCAAGAAACTTTTGGGTAAATATCAACATAAATCTCAGGCAGGAACGA	5040	6061	TCTAATACATATGATCAATACACTTTTGTGTTTAAAAAGAAATTAATACTTATTTCAAAC	6120
Qy	5041	GGAATACACATCATCGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCAACTCAAC	5100	6121	ATCGATCACATTTTACTTTTGTGTTTCCATATGACTACATTTATATAGGCTCACACTTTGT	6180
Db	5041	GGAATACACATCATCGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCAACTCAAC	5100	6121	ATCGATCACATTTTACTTTTGTGTTTCCATATGACTACATTTATATAGGCTCACACTTTGT	6180
Qy	5101	AGGCCATCTCTCATTTCTAAGTACTCAGCACTAGCTATTTCTTGATGATTTCTTTTGGTT	5160	6241	TTTCCACGATGTTCTCTGGTGGGTGGGTCCATCTTTTGGGACCACTGTGCGGTAGAGGAT	6300
Db	5101	AGGCCATCTCTCATTTCTAAGTACTCAGCACTAGCTATTTCTTGATGATTTCTTTTGGTT	5160	6241	TTTCCACGATGTTCTCTGGTGGGTGGGTCCATCTTTTGGGACCACTGTGCGGTAGAGGAT	6300
Qy	5161	GGGGTGATCAATTTGGTGTGTCATGCGAGTTATTCGTGAGGAAGATCAAAACATCAGCT	5220	6301	CTTGAAACGATAGCTTTCTTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTTTT	6360
Db	5161	GGGGTGATCAATTTGGTGTGTCATGCGAGTTATTCGTGAGGAAGATCAAAACATCAGCT	5220	6301	CTTGAAACGATAGCTTTCTTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTTTT	6360
Qy	5221	ATATCCAAAGGCTAATTTTGGGCTCAAGGAAGGTATGGTTATAAACATATCTTTTG	5280	6421	ATATTACCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTTCTGAGACTGTATCTTTGA	6480
Db	5221	ATATCCAAAGGCTAATTTTGGGCTCAAGGAAGGTATGGTTATAAACATATCTTTTG	5280	6421	ATATTACCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTTCTGAGACTGTATCTTTGA	6480
Qy	5281	ATCTTTTAAAGATCTTCAAAAGTGTGAGTATGTTTATTTGGTGGCTTTCTGGTGATTTTA	5340	6481	TATTTCTTGAGTACGAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCACATC	6540
Db	5281	ATCTTTTAAAGATCTTCAAAAGTGTGAGTATGTTTATTTGGTGGCTTTCTGGTGATTTTA	5340	6481	TATTTCTTGAGTACGAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCACATC	6540
Qy	5341	TGTTTTATPAGAATTTGGTCTTATATATTTGGTCTATATATAGAGTGTGGGTGATATGAT	5400	6541	AATCCACTTGTCTTGAAGACGTGGTTGGAACGTCTTTTTCACAGATGTTCTCGTGG	6600
Db	5341	TGTTTTATPAGAATTTGGTCTTATATATTTGGTCTATATATAGAGTGTGGGTGATATGAT	5400	6541	AATCCACTTGTCTTGAAGACGTGGTTGGAACGTCTTTTTCACAGATGTTCTCGTGG	6600
Qy	5401	GAAATCAAGACTGATTTGGAACTTTTCTGTGTTCAATGAATATCATCGAATCT	5460	6601	GTGGGGTCCATCTTTTGGGACCACTGTGCGTAGAGGCATCTTTGAAACGATAGCTTTCCCT	6660
Db	5401	GAAATCAAGACTGATTTGGAACTTTTCTGTGTTCAATGAATATCATCGAATCT	5460	6601	GTGGGGTCCATCTTTTGGGACCACTGTGCGTAGAGGCATCTTTGAAACGATAGCTTTCCCT	6660
Qy	5461	CAATTTCTTGGAGACCATATATGAGACATGAGACATCTATAGAACATATATGTAATGTA	5520	6661	TATCCCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTCTACTGTCTCTTCGATGAAGT	6720
Db	5461	CAATTTCTTGGAGACCATATATGAGACATGAGACATCTATAGAACATATATGTAATGTA	5520	6661	TATCCCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTCTACTGTCTCTTCGATGAAGT	6720
Qy	5521	TATTTAAACGTACTTAAGTCGAATTTTATGACCAAGTAAATAAATATGCGGAATGTACA	5580	6721	GACAGATAGCTGGGCAATGGAAATCCGAGAGGTTTCCGATATACCTTCTGTAAGAAG	6780
Db	5521	TATTTAAACGTACTTAAGTCGAATTTTATGACCAAGTAAATAAATATGCGGAATGTACA	5580	6721	GACAGATAGCTGGGCAATGGAAATCCGAGAGGTTTCCGATATATACCTTCTGTAAGAAG	6780
Qy	5581	TGCTAAATATCGAGTTTAAACTATTTTCCAAATATAAACAATATTTTCTCTTTCCGTC	5640	6781	TCCTCAATAGCCCTCTGCTCTGAGACTGTATCTTTGATATCTTTGAGTATCTTTGGAGTACGAGAG	6840
Db	5581	TGCTAAATATCGAGTTTAAACTATTTTCCAAATATAAACAATATTTTCTCTTTCCGTC	5640	6781	TCCTCAATAGCCCTCTGCTCTGAGACTGTATCTTTGATATCTTTGAGTATCTTTGGAGTACGAGAG	6840
Qy	5641	CTTATATATCTTATTTCTGATTTCTTATTTTCTTTTAAATTTCTTTTCTTTTCCCAA	5700	6841	TGTCGTGCTCCACCATGTTGGGGATCTAGATATCATCAATCCATCCACTTCTTTTGAAGACG	6900
Db	5641	CTTATATATCTTATTTCTGATTTCTTATTTTCTTTTAAATTTCTTTTCTTTTCCCAA	5700	6841	TGTCGTGCTCCACCATGTTGGGGATCTAGATATCATCAATCCATCCACTTCTTTTGAAGACG	6900
Qy	5701	GACACAAAAAATAAATAACAGAAACGAAAAAAGAGATTTTAAAAATTCATTAACCCAC	5760			
Db	5701	GACACAAAAAATAAATAACAGAAACGAAAAAAGAGATTTTAAAAATTCATTAACCCAC	5760			
Qy	5761	GAGAAATTATGCACCTAAATTCAGACTAATCCCCCAAAATTTTCAGAAAATTTATGTAATTTTTG	5820			

6901 TGGTTGGAAAGCTCTCTCTTTTCCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGGAC 6960  
6901 TGGTTGGAAAGCTCTCTCTTTTCCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGGAC 6960  
6961 CACTGTCGGTAGAGGCATCTTGAACGATAGCCCTTTCTTTTATCGCAATGATGGCAATTTGT 7020  
6961 CACTGTCGGTAGAGGCATCTTGAACGATAGCCCTTTCTTTTATCGCAATGATGGCAATTTGT 7020  
7021 AGAAGCCATCTTCTCTTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGA 7080  
7021 AGAAGCCATCTTCTCTTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGA 7080  
7081 ATCCGAGGAGGTTTCCGATATATACCTTTGTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTT 7140  
7081 ATCCGAGGAGGTTTCCGATATATACCTTTGTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTT 7140  
7141 CTGAGACTGATCTTTGATATCTTGGAGTGACGAGAGTGTCGTCTCCACCATGTTGG 7200  
7141 CTGAGACTGATCTTTGATATCTTGGAGTGACGAGAGTGTCGTCTCCACCATGTTGG 7200  
7201 GGATCTAGATATACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTCTCTTTT 7260  
7201 GGATCTAGATATACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTCTCTTTT 7260  
7261 CCAAGATGTTCTCTGTTGGGTGGGGTCCATCTTTTGGGACCACTGTCGGTAGAGGCATCTT 7320  
7261 CCAAGATGTTCTCTGTTGGGTGGGGTCCATCTTTTGGGACCACTGTCGGTAGAGGCATCTT 7320  
7321 GAACGATAGCTTCTCTTTATCGCAATGATGCAATTTGAGAACCATCTCTCTTTCTA 7380  
7321 GAACGATAGCTTCTCTTTATCGCAATGATGCAATTTGAGAACCATCTCTCTTTCTA 7380  
7381 CTGTCTCTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATA 7440  
7381 CTGTCTCTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATA 7440  
7441 TTACCTCTTTGTTGAAAAGTCTCAATAGCCCTCTGTTCTCTCGAGACTGTATCTTTGATAT 7500  
7441 TTACCTCTTTGTTGAAAAGTCTCAATAGCCCTCTGTTCTCTCGAGACTGTATCTTTGATAT 7500  
7501 TCTTGGAGTAGACGAGAGTGTCGTCTCCACCATGTTGGGATCCACTAGTTCAGAGCG 7560  
7501 TCTTGGAGTAGACGAGAGTGTCGTCTCCACCATGTTGGGATCCACTAGTTCAGAGCG 7560  
7561 GCCGCCACCGCGGTGGAGCT 7580  
7561 GCCGCCACCGCGGTGGAGCT 7580

RESULT 2  
US-10-650-249-18  
; Sequence 18, Application US/10650249  
; Publication No. US20040045055A1  
; GENERAL INFORMATION:  
; APPLICANT: Neff, Michael M  
; TITLE OF INVENTION: THE GENE FOR A DOF TRANSCRIPTION FACTOR CAPABLE OF ALTERING  
; TITLE OF INVENTION: THE SIZE AND STATURE OF A PLANT  
; FILE REFERENCE: WSHU 2064.1  
; CURRENT APPLICATION NUMBER: US/10/650,249  
; CURRENT FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US 60/406,657  
; PRIOR FILING DATE: 2002-08-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 18  
; LENGTH: 1746  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-650-249-18

Query Match 22.8%; Score 1730; DB 7; Length 1746;  
Best Local Similarity 99.8%; Pred. No. 3.9e-305;

Matches 1743; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 4436 AACAGAGAGAGCAATCCAGATCGAAATCTACGGTCGTGGTCTCGACTGATTAATCTACT 4495  
DB 1 AACAGAGAGAGCAATCCAGATCGAAATCTACGGTCGTGGTCTCGACTGATTAATCTACT 60  
QY 4496 AGTACTCTCATCACTTACTTCTCGCCCAAGTTACTCAAAACCCCTAGCAAGTTTCATAGCTAC 4555  
DB 61 AGTACTCTCATCACTTACTTCTCGCCCAAGTTACTCAAAACCCCTAGCAAGTTTCATAGCTAC 120  
QY 4556 GGTCAAAATCCCGGAGTTTAAATTCCAACTTGGCCCATCTTGGCCCTCTCTCCAAAGCCCTTGA 4615  
DB 121 GGTCAAAATCCCGGAGTTTAAATTCCAACTTGGCCCATCTTGGCCCTCTCTCCAAAGCCCTTGA 180  
QY 4616 GATTAACAATCAAGCAACACATCGGATTTAGATTTTGGTGGAACTCAATAAAGCAACATGATA 4675  
DB 181 GATTAACAATCAAGCAACACATCGGATTTAGATTTTGGTGGAACTCAATAAAGCAACATGATA 240  
QY 4676 AGTGTGTAGCTTCTAGTGGTGGGATCTTGGATGATGGAGAAATACCTCCATCAACAACAA 4735  
DB 241 AGTGTGTAGCTTCTAGTGGTGGGATCTTGGATGATGGAGAAATACCTCCATCAACAACAA 300  
QY 4736 GCTCAGCAATTCCTTTCTTGATCAACACTACCGGATTTGGTGAATCTTTCAAAACGCTTGA 4795  
DB 301 GCTCAGCAATTCCTTTCTTGATCAACACTACCGGATTTGGTGAATCTTTCAAAACGCTTGA 360  
QY 4796 TATCCATTTACTAGAAAGTAAAGGAGGTGTAAATCAAGGTGATTTCAACAGAAAGTGTAGT 4855  
DB 361 TATCCATTTACTAGAAAGTAAAGGAGGTGTAAATCAAGGTGATTTCAACAGAAAGTGTAGT 420  
QY 4856 GATTAATTCGAATCAGCTAATGTTTAAAGCCCTTGATGGATTTTCTTCAGGCGGGGTAGC 4915  
DB 421 GATTAATTCGAATCAGCTAATGTTTAAAGCCCTTGATGGATTTTCTTCAGGCGGGGTAGC 480  
QY 4916 GCCAGCAACCAAGAAATGTGAAGCGGAAGAAATGATCAGGATCGGGGTAGGATGGG 4975  
DB 481 GCCAGCAACCAAGAAATGTGAAGCGGAAGAAATGATCAGGATCGGGGTAGGATGGG 540  
QY 4976 GATGAGTGAATTAATCAAGAAATCTTTTGGGTAAATCAACATAACTCAGGCAGG 5035  
DB 541 GATGAGTGAATTAATCAAGAAATCTTTTGGGTAAATCAACATAACTCAGGCAGG 600  
QY 5036 AACGAGGAATACATCATGCGGAGGTAAAGTCTTGGACCGGTTTCACTTCCAAACAAC 5095  
DB 601 AACGAGGAATACATCATGCGGAGGTAAAGTCTTGGACCGGTTTCACTTCCAAACAAC 660  
QY 5096 TCAACAGGCCATCTCTCAATTC--TAAGTACTCAGCACTAGCTATTTCTTGATGATTTCTTT 5152  
DB 661 TCAACAGGCCATCTCTCAATTC--TAAGTACTCAGCACTAGCTATTTCTTGATGATTTCTTT 720  
QY 5153 TGTGGTTGGGGTGATCATTTGGTGTCTGTATGTCGAGTTATTTGTCGAGGAAGATCAAAAC 5212  
DB 721 TGTGGTTGGGGTGATCATTTGGTGTCTGTATGTCGAGTTATTTGTCGAGGAAGATCAAAAC 780  
QY 5213 ATGCACTATATCCAAAGCTTAATTTGAGGCTCAAGGCTCAAGGAAAGGTATGGTTATAAACTA 5272  
DB 781 ATGCACTATATCCAAAGCTTAATTTGAGGCTCAAGGAAAGGTATGGTTATAAACTA 840  
QY 5273 TCTTTTGTATCTTTTAAAGATCTTCAAGGTGAGTATGTTTATTTGGTGGCTTCTGGT 5332  
DB 841 TCTTTTGTATCTTTTAAAGATCTTCAAGGTGAGTATGTTTATTTGGTGGCTTCTGGT 900  
QY 5333 GATATTTATGTTTATTAAGATTTTGGTCTTATATATGCTATATATAGAGGTGGGGTG 5392  
DB 901 GATATTTATGTTTATTAAGATTTTGGTCTTATATATGCTATATATAGAGGTGGGGTG 960  
QY 5393 ATATGATGAATTCAGAGTTGATGTTGGAACCTTTTGTGTGTTCATTTGAATTAATCAT 5452  
DB 961 ATATGATGAATTCAGAGTTGATGTTGGAACCTTTTGTGTGTTCATTTGAATTAATCAT 1020  
QY 5453 CGAATTTCAATTTCTGGAGCCCATTTATGAGCAATTTAGACATCTATAGAACATATAT 5512  
DB 1021 CGAATTTCAATTTCTGGAGCCCATTTATGAGCAATTTAGACATCTATAGAACATATAT 1080



5513 GTAATGTATATTAACGCTACTTAAGTCGAATTTTATGACCAAGTAAATATAATATATGCGG 5572  
1081 GTAATGTATATTAACGCTACTTAAGTCGAATTTTATGACCAAGTAAATATAATATATGCGG 1140  
5573 AATGTACATGCTAAATATCGAGTTTAAACTATTTTCCAAATATACAACTATTTTCTCTT 5632  
1141 AATGTACATGCTAAATATCGAGTTTAAACTATTTTCCAAATATACAACTATTTTCTCTT 1200  
5633 TCGTCCAACTTATATACCTTATCTGATCTTATTTCTTCTTTTAAATTCCTTTTCC 5692  
1201 TCGTCCAACTTATATACCTTATCTGATCTTATTTCTTCTTTTAAATTCCTTTTCC 1260  
5693 TTTCCCAAGACAAATAAATAATATACAGAAACGAAAGAGATTTTAAATAATCA 5752  
1261 TTTCCCAAGACAAATAAATAATATACAGAAACGAAAGAGATTTTAAATAATCA 1320  
5753 TAAACCCAGAGAAATATGACCTAAATTCAGACTAATCCCCCAAAATTTTCAGAAATTTATG 5812  
1321 TAAACCCAGAGAAATATGACCTAAATTCAGACTAATCCCCCAAAATTTTCAGAAATTTATG 1380  
5813 TATTTTTCGGAATTAATATGTTTCAAAATCAATATGAGCCCAACTAACTAATTTGAAAGA 5872  
1381 TATTTTTCGGAATTAATATGTTTCAAAATCAATATGAGCCCAACTAACTAATTTGAAAGA 1440  
5873 CAATGGAATGACTGAAACCAATGCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 5932  
1441 CAATGGAATGACTGAAACCAATGCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCAAGTCTCA 1500  
5933 ACCAATAGACTATCATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5992  
1501 ACCAATAGACTATCATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
5993 GATATGCTCATTTATCTGGATATAAGATGGCGCTTTTAACTCTCTTTGCAATTTTGTAT 6052  
1561 GATATGCTCATTTATCTGGATATAAGATGGCGCTTTTAACTCTCTTTGCAATTTTGTAT 1620  
6053 ATCTTTCTTCTAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6112  
1621 ATCTTTCTTCTAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
6113 TTTCAACATCGATCATATTTTACTTTTGTTCATATGATGATGATGATGATGATGATGATGATGAT 6172  
1681 TTTCAACATCGATCATATTTTACTTTTGTTCATATGATGATGATGATGATGATGATGATGATGAT 1740  
6173 ACTTTT 6178  
1741 ACTTTT 1746

RESULT 3  
US-10-033-190-3/c  
; Sequence 3, Application US/10033190  
; Publication No. US20020133848A1  
; GENERAL INFORMATION:  
; APPLICANT: Exelixis Plant Sciences, Inc.  
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF AN ANTHOCYANIN MUTANT (ANT  
; TITLE OF INVENTION: TOMATO  
; FILE REFERENCE: EP01-002C  
; CURRENT APPLICATION NUMBER: US/10/033,190  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: US 60/244,685  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 10078  
; TYPE: DNA  
; ORGANISM: pag3202  
US-10-033-190-3

Query Match 18.5%; Score 1402.4; DB 5; Length 10078;  
Best Local Similarity 99.9%; Pred. No. 4.9e-245;

Matches 1403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	6177	TTGTTTCGGATCTAGATATACATCAATCCACTTTCCTTTGAAGACGTGGTTGGAAACGCTCT	6236
DB	4251	TAGTTTCGGATCTAGATATACATCAATCCACTTTCCTTTGAAGACGTGGTTGGAAACGCTCT	4192
QY	6237	TCTTTTTCACGATGTTCTCGTGGGTGGGGTCCATCTTTTGGGACCACTGTCCGTAGAG	6296
DB	4191	TCTTTTTCACGATGTTCTCGTGGGTGGGGTCCATCTTTTGGGACCACTGTCCGTAGAG	4132
QY	6297	GCATCTTGAAACGATAGCCTTTTCCCTTTATCGCAATGATGGCAATTTGTAGAACCATCTTCC	6356
DB	4131	GCATCTTGAAACGATAGCCTTTTCCCTTTATCGCAATGATGGCAATTTGTAGAACCATCTTCC	4072
QY	6357	TTTTCTACTGTCCTTTTCGATGAAGTGCAGATAGCTGGGCAATGGAATCCGAGGAGGTTT	6416
DB	4071	TTTTCTACTGTCCTTTTCGATGAAGTGCAGATAGCTGGGCAATGGAATCCGAGGAGGTTT	4012
QY	6417	CCCGATATTACCTTTTGTGAAAGTCTCAATAGCCCTCTCGTCTTCTGAGACTGTATCT	6476
DB	4011	CCCGATATTACCTTTTGTGAAAGTCTCAATAGCCCTCTCGTCTTCTGAGACTGTATCT	3952
QY	6477	TTGATATTCTTTGGAGTAGACGAGAGTGTCTCCACCATTTGGGGATCTAGATATCA	6536
DB	3951	TTGATATTCTTTGGAGTAGACGAGAGTGTCTCTCCACCATTTGGGGATCTAGATATCA	3892
QY	6537	CATCAATCCACTTGTGTTGAAAGAGTGGTTCGAAAGTCTTCTTTTCCAGATGTTCTCTC	6596
DB	3891	CATCAATCCACTTGTGTTGAAAGAGTGGTTCGAAAGTCTTCTTTTCCAGATGTTCTCTC	3832
QY	6597	GTGGGTGGGGGTCCATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCCTT	6656
DB	3831	GTGGGTGGGGGTCCATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCCTT	3772
QY	6657	CCTTTATCGCAATGATGGCAATTTGAGAACCATCTCTCTTTCTACTGTCTTTCGATG	6716
DB	3771	CCTTTATCGCAATGATGGCAATTTGAGAACCATCTCTCTTTCTACTGTCTTTCGATG	3712
QY	6717	AAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGA	6776
DB	3711	AAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGA	3652
QY	6777	AAAGTCTCAATAGCCCTCTGGTCTTCTCAGACTGTATCTTTGATATTTCTGGAGTAGACG	6836
DB	3651	AAAGTCTCAATAGCCCTCTGGTCTTCTCAGACTGTATCTTTGATATTTCTGGAGTAGACG	3592
QY	6837	AGAGTGTCTGCTCCACCATGTTGGGGATCTAGATATCACATCAATCCACTTTGCTTTGAA	6896
DB	3591	AGAGTGTCTGCTCCACCATGTTGGGGATCTAGATATCACATCAATCCACTTTGCTTTGAA	3532
QY	6897	GACGTGGTTGGAACGTCTTCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTTG	6956
DB	3531	GACGTGGTTGGAACGTCTTCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTTG	3472
QY	6957	GGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCTTTTCTTTATCGCAATGATGGCAT	7016
DB	3471	GGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCTTTTCTTTATCGCAATGATGGCAT	3412
QY	7017	TTGTAGAAAGCCATCTTCTCTTTCTACTGTCTTTTCGATGAAGTGCAGATAGCTGGGCAA	7076
DB	3411	TTGTAGAAAGCCATCTTCTCTTTCTACTGTCTTTTCGATGAAGTGCAGATAGCTGGGCAA	3352
QY	7077	TGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGAAAGTCTCAATAGCCCTCTGG	7136
DB	3351	TGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGAAAGTCTCAATAGCCCTCTGG	3292
QY	7137	TCCTCTTGAGACTGTATCTTTGATATTTCTTGAGTAGAGAGTGTCTGCTCCACCATG	7196
DB	3291	TCCTCTTGAGACTGTATCTTTGATATTTCTTGAGTAGAGAGTGTCTGCTCCACCATG	3232
QY	7197	TTGGGGATCTAGATATCAATCAATCCACTTCTGCTTTGAGAGCGTGGTTGGAAACGCTTCT	7256
DB	3231	TTGGGGATCTAGATATCAATCAATCCACTTCTGCTTTGAGAGCGTGGTTGGAAACGCTTCT	3172

QY 7257 TTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGGTAGAGCA 7316  
DB 3171 TTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGGTAGAGCA 3112  
QY 7317 TCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCCTT 7376  
DB 3111 TCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCCTT 3052  
QY 7377 TCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCC 7436  
DB 3051 TCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCC 2992  
QY 7437 GATATACCTTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTG 7496  
DB 2991 GATATACCTTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTG 2932  
QY 7497 ATATTCTTGGAGTAGACGAGAGTGTGCTCCACCATGTTGGGATCCACTAGTCTAG 7556  
DB 2931 ATATTCTTGGAGTAGACGAGAGTGTGCTCCACCATGTTGGGATCCACTAGTCTAG 2872  
QY 7557 AGCGGCCGCCACCGGGTGGAGCT 7580  
DB 2871 AGCGGCCGCCACCGGGTGGAGCT 2848

## RESULT 4

US-09-522-334-1/c

; Sequence 1, Application US/09522334

; Patent No. US20020157130A1

; GENERAL INFORMATION:

; APPLICANT: Wagner, Ry

; APPLICANT: Mathews, Helena

; APPLICANT: Liu, Xing Liang

; APPLICANT: Waggoner, Wency J.

; TITLE OF INVENTION: TRAIT-ASSOCIATED GENE IDENTIFICATION

; TITLE OF INVENTION: METHOD

; FILE REFERENCE: 4257-0018.30

; CURRENT APPLICATION NUMBER: US/09/522,334

; CURRENT FILING DATE: 2000-03-09

; PRIOR APPLICATION NUMBER: US 60/124,232

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 1361

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: modified enhancer

US-09-522-334-1

Query Match 17.9%; Score 1360; DB 3; Length 1361;  
Best Local Similarity 100.0%; Pred. No. 1.1e-237; Mismatches 0; Gaps 0;  
Matches 1360; Conservative 0; Indels 0

QY 6179 GTTTCGGATCTAGATATCACATCAATCCACTTGTGAAACGCTGTTGAAACGCTCTTC 6238  
DB 1360 GTTTCGGATCTAGATATCACATCAATCCACTTGTGAAACGCTGTTGAAACGCTCTTC 1301  
QY 6239 TTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGACCACTGTGGTAGAGGC 6298  
DB 1300 TTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGACCACTGTGGTAGAGGC 1241  
QY 6299 ATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTT 6358  
DB 1240 ATCTTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTT 1181  
QY 6359 TTCTACTGTCTTTTCGATGAAGTGACAGTACCTGGCAATGGAATCCGAGGAGTTTCC 6418  
DB 1180 TTCTACTGTCTTTTCGATGAAGTGACAGTACCTGGCAATGGAATCCGAGGAGTTTCC 1121  
QY 6419 CGATATTACCCCTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTT 6478

DB 1120 CGATATTACCCCTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTT 1061  
QY 6479 GATATTTCTTGGAGTAGACGAGAGTGTCTGTCTCCACCATGTTGGGATCTAGATATCACA 6538  
DB 1060 GATATTTCTTGGAGTAGACGAGAGTGTCTGTCTCCACCATGTTGGGATCTAGATATCACA 1001  
QY 6539 TCAATCCACTTGTCTTTGAAAGCGTGGTGGAAAGTCTTCTTTTCCAGATGTTCTCTGT 6598  
DB 1000 TCAATCCACTTGTCTTTGAAAGCGTGGTGGAAAGTCTTCTTTTCCAGATGTTCTCTGT 941  
QY 6599 GGGTGGGGTCCATCTTTGGGACCACTGTCTGTAGAGGAGTCTTCTTGAACAGATAGCCTTTCC 6658  
DB 940 GGGTGGGGTCCATCTTTGGGACCACTGTCTGTAGAGGAGTCTTCTTGAACAGATAGCCTTTCC 881  
QY 6659 TTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTCTACTGTCTTTTCGATGAA 6718  
DB 880 TTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTCTACTGTCTTTTCGATGAA 821  
QY 6719 GTGACAGATAGTGGGCAATGGAATCCGAGGAGTTTCCGATATATACCTTTTGTGAAA 6778  
DB 820 GTGACAGATAGTGGGCAATGGAATCCGAGGAGTTTCCGATATATACCTTTTGTGAAA 761  
QY 6779 AGTCTCAATAGCCCTCTGTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAG 6838  
DB 760 AGTCTCAATAGCCCTCTGTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAG 701  
QY 6839 AGTGTCTGTCTCCACCATGTTGGGATCTAGATATCACATCAATCCACTTGTCTTTGAAAGA 6898  
DB 700 AGTGTCTGTCTCCACCATGTTGGGATCTAGATATCACATCAATCCACTTGTCTTTGAAAGA 641  
QY 6899 CGTGGTTGGAACGCTCTTCTTTTCCAGATGTTCTCTGTGGTGGGGTCCATCTTTGGG 6958  
DB 640 CGTGGTTGGAACGCTCTTCTTTTCCAGATGTTCTCTGTGGTGGGGTCCATCTTTGGG 581  
QY 6959 ACCACTGTCTGTAGAGGCACTTGAACAGATAGCCTTCTTATCGCAATGATGCAATTT 7018  
DB 580 ACCACTGTCTGTAGAGGCACTTGAACAGATAGCCTTCTTATCGCAATGATGCAATTT 521  
QY 7019 GTAGAAGCCATCTTCTTCTTCTACTGTCTTCTGATGAAGTGACAGATAGCTGGGCAATG 7078  
DB 520 GTAGAAGCCATCTTCTTCTTCTACTGTCTTCTGATGAAGTGACAGATAGCTGGGCAATG 461  
QY 7079 GAATCCGAGGAGTTTCCGATATATACCTTTGTTGAAAAGTCTCAATAGCCCTCTGTGTC 7138  
DB 460 GAATCCGAGGAGTTTCCGATATATACCTTTGTTGAAAAGTCTCAATAGCCCTCTGTGTC 401  
QY 7139 TTCTGAGACTGTATCTTTGATATTTCTGAGTAGACGAGTGTCTGCTCCACCATGTT 7198  
DB 400 TTCTGAGACTGTATCTTTGATATTTCTGAGTAGACGAGTGTCTGCTCCACCATGTT 341  
QY 7199 GGGGATCTAGATATCACATCAATCCACTTGTGAAACGCTGTTGAAACGCTCTTCTTT 7258  
DB 340 GGGGATCTAGATATCACATCAATCCACTTGTGAAACGCTGTTGAAACGCTCTTCTTT 281  
QY 7259 TTCCACGATGTTCTCTGTGGGTGGGGTCCATCTTTGGGACCACTGTGGTAGAGGCATC 7318  
DB 280 TTCCACGATGTTCTCTGTGGGTGGGGTCCATCTTTGGGACCACTGTGGTAGAGGCATC 221  
QY 7319 TTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 7378  
DB 220 TTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 161  
QY 7379 TACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTGTCTTCCC 7438  
DB 160 TACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTGTCTTCCC 101  
QY 7439 TATTACCCCTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGAT 7498  
DB 100 TATTACCCCTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGAT 41  
QY 7499 ATTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTG 7538





```
QY 5237 TTTGAGGCTCAAGGAAAGGATGTTATATAAACTATCTTTTGTGATCTTTTAAAGATCT 5296
Db 121 TTTGAGGCTCAAGGAAAGGATGTTATATAAACTATCTTTTGTGATCTTTTAAAGATCT 180
QY 5297 TCAAAAGTGTAGTATGTTTATTTGTTGGTCTTCTGGTGATATTATGTTTATTAAGAAATTT 5356
Db 181 TCAAAAGTGTAGTATGTTTATTTGTTGGTCTTCTGGTGATATTATGTTTATTAAGAAATTT 240
QY 5357 GGTCTTATATATTTGGCTATATATAGAGGTGTGGGTGATGATGTAGTAATTCAGAGTTGAT 5416
Db 241 GGTCTTATATATTTGGCTATATATAGAGGTGTGGGTGATGATGTAGTAATTCAGAGTTGAT 300
QY 5417 GTTGGAAACTTTTGTGCTGTTCAATGTAATCATCGAATTCCTCAATTTCTTGGAGACC 5476
Db 301 GTTGGAAACTTTTGTGCTGTTCAATGTAATCATCGAATTCCTCAATTTCTTGGAGACC 360
QY 5477 CATTATGAGACATTCAGACATCTATAGAACATATATATGTAATGTAATTAATAACGCTACTTAA 5536
Db 361 CATTATGAGACATTCAGACATCTATAGAACATATATGTAATGTAATTAATAACGCTACTTAA 420
QY 5537 GTCGAAATTTTATGACCAAGTAAATTAATATGCGGAATGACATGCTTAATATCGAGTTT 5596
Db 421 GTCGAAATTTTATGACCAAGTAAATTAATATGCGGAATGACATGCTTAATATCGAGTTT 480
QY 5597 AAACATATTTTCCNATATAACAATCTATTTCTCTTTCGTCGAATCTATCTTAT 5656
Db 481 AAACATATTTTCCNATATAACAATCTATTTCTCTTTCGTCGAATCTATCTTAT 540
QY 5657 CTGATCTTATTTCTCTTTTAACTTCTTTTCTTTTCTTTCCTTCCAGACACAAAAA 5716
Db 541 CTGATCTTATTTCTCTTTTAACTTCTTTTCTTTTCTTTCCTTCCAGACACAAAAA 600
QY 5717 AATCAGAAACGAAAAAGAGATTTTAAAAATTCATACCCACGAGAAATATGCACCTA 5776
Db 601 AATCAGAAACGAAAAAGAGATTTTAAAAATTCATACCCACGAGAAATATGCACCTA 660
QY 5777 AATTCAGACTAATCCCCCAATTTTCAGAAATTTATGATATTTTGGGATTTAATTTGTT 5836
Db 661 AATTCAGACTAATCCCCCAATTTTCAGAAATTTATGATATTTTGGGATTTAATTTGTT 720
QY 5837 TCAGAACTAATGCGCAACTAACTAACTTGAAGACAAATGAAATGAACTGAAACCATGCA 5896
Db 721 TCAGAACTAATGCGCAACTAACTAACTTGAAGACAAATGAAATGAACTGAAACCATGCA 780
QY 5897 TAATCTCTCAAGTCTCAACCTATGAAGAAATCATGTAACTAAGACTATCATGATTA 5956
Db 781 TAATCTCTCAAGTCTCAACCTATGAAGAAATCATGTAACTAAGACTATCATGATTA 840
QY 5957 GTTAATGCATGATCTATATGATATTTCTTTGAACATAGATATGTCATTTATCTGGATATA 6016
Db 841 GTTAATGCATGATCTATATGATATTTCTTTGAACATAGATATGTCATTTATCTGGATATA 900
QY 6017 AGATGGCGTTTAACTACTTTTGAATTTTGTATATCTTTCTTAATACATATGATC 6076
Db 901 AGATGGCGTTTAACTACTTTTGAATTTTGTATATCTTTCTTAATACATATGATC 960
QY 6077 AATACACTTTTGTTTTAAAGAAATTAATAAATTAATTTCAACATCGATCACAATTTTA 6136
Db 961 AATACACTTTTGTTTTAAAGAAATTAATAAATTAATTTCAACATCGATCACAATTTTA 1020
QY 6137 CTTTGTGTTCCATATGATCAATTTATAGGCTCACATTTT 6178
Db 1021 CTTTGTGTTCCATATGATCAATTTATAGGCTCACATTTT 1062
```

## RESULT 7

```
US-10-650-249-19
; Sequence 19, Application US/10650249
; Publication No. US20040045055A1
; GENERAL INFORMATION:
; APPLICANT: Neff, Michael M
; TITLE OF INVENTION: THE GENE FOR A DOF TRANSCRIPTION FACTOR CAPABLE OF ALTERING
; TITLE OF INVENTION: THE SIZE AND STATURE OF A PLANT
```

```
; FILE REFERENCE: WSHU 2064.1
; CURRENT APPLICATION NUMBER: US/10/650.249
; PRIOR FILING DATE: 2003-08-28
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1058
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-650-249-19
```

```
Query Match 14.0%; Score 1058; DB 7; Length 1058;
Best Local Similarity 100.0%; Pred. No. 1.2e-182; Indels 0; Gaps 0;
Matches 1058; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3045 CAACCAAGAACGATGACGTATATGATTCGACCTTGCAAAAAATAGCAAAACAAATACCTGTT 3104
Db 1 CAACCAAGAACGATGACGTATATGATTCGACCTTGCAAAAAATAGCAAAACAAATACCTGTT 60
QY 3105 CAAATCGACACTTAATTCCAAAAAAGGTTAGTAATAAGTAAGCAAGGCTTTTATTAAGAAA 3164
Db 61 CAAATCGACACTTAATTCCAAAAAAGGTTAGTAATAAGTAAGCAAGGCTTTTATTAAGAAA 120
QY 3165 ACAAAGAAATAAAGAGCCTTAAGAGATGATGAAATTTGAAGAGAAAAAGAGCATTTG 3224
Db 121 ACAAAGAAATAAAGAGCCTTAAGAGATGATGAAATTTGAAGAGAAAAAGAGCATTTG 180
QY 3225 TTATAGAAAAAGAAAAAGAGAGAGTAAAGAGAAATTAAGAAACAAATTAATAACAA 3284
Db 181 TTATAGAAAAAGAAAAAGAGAGAGTAAAGAGAAATTAAGAAACAAATTAATAACAA 240
QY 3285 AGGAAACTTCAATTTCTTCTTTATCCCATTCAGCTCTCTCTCTCTCTCTCTCTCTCTCT 3344
Db 241 AGGAAACTTCAATTTCTTCTTTATCCCATTCAGCTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 3345 CTCTCTCTCTAGATCAATTTCTTCTTATGATGATGATTAATCCACCATATCTCGGACC 3404
Db 301 CTCTCTCTCTAGATCAATTTCTTCTTATGATGATGATTAATCCACCATATCTCGGACC 360
QY 3405 TCTTACCTTAAAGAGATACAAAGTAAAGATTTCAAAGATGGTTTTCTCATCTCTCCAGTG 3464
Db 361 TCTTACCTTAAAGAGATACAAAGTAAAGATTTCAAAGATGGTTTTCTCATCTCTCCAGTG 420
QY 3465 AATCAGTTTCCGATTTCCCAAAATTTGGCAGCAGGTAAAAATCAGTTTATGATATTTGCTAGAT 3524
Db 421 AATCAGTTTCCGATTTCCCAAAATTTGGCAGCAGGTAAAAATCAGTTTATGATATTTGCTAGAT 480
QY 3525 GTTTCGATTCGTTCTCTTTTCTTCCCAAGCTCGATCAAGATTTATGAAAAATTTGATGAGA 3584
Db 481 GTTTCGATTCGTTCTCTTTTCTTCCCAAGCTCGATCAAGATTTATGAAAAATTTGATGAGA 540
QY 3585 TTTTGTTCGACAAAAATTCCTAGCTATTTGTGACGCGCATATATATTACTTATGATATTTTC 3644
Db 541 TTTTGTTCGACAAAAATTCCTAGCTATTTGTGACGCGCATATATATTACTTATGATATTTTC 600
QY 3645 TTAGTTGATTAACCCCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3704
Db 601 TTAGTTGATTAACCCCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
QY 3705 GATTTCAAATTTGGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3764
Db 661 GATTTCAAATTTGGTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
QY 3765 TAAAAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3824
Db 721 TAAAAACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
QY 3825 AATTTCTCTTTGTTAGATTTTACACCGCATGGAATTAATCACTTCAAAAATAAAAAA 3884
Db 781 AATTTCTCTTTGTTAGATTTTACACCGCATGGAATTAATCACTTCAAAAATAAAAAA 840
```

```
QY 3885 GTTTAAAGTACTATGACTTTAACTGAGTTATTTATCCATTTCTTTTTCAGCTTGT 3944
Db 841 GTTTAAAGTACTATGACTTTAACTGAGTTATTTATCCATTTCTTTTTCAGCTTGT 900
QY 3945 TGAAAACTAATAATTAATCTGCAATTCCTGCAAGTAGTCACAAATTTTATCTATTTTC 4004
Db 901 TGAAAACTAATAATTAATCTGCAATTCCTGCAAGTAGTCACAAATTTTATCTATTTTC 960
QY 4005 TTTTGTCTCCAGCAATGTTTCAAACTCGAATCTTCTGTTTAAAGTGTCTTCTGCTTTA 4064
Db 961 TTTTGTCTCCAGCAATGTTTCAAACTCGAATCTTCTGTTTAAAGTGTCTTCTGCTTTA 1020
QY 4065 TTATAAACTCGAACTAATTAAGTACAAATTAATGTTAAT 4102
Db 1021 TTATAAACTCGAACTAATTAAGTACAAATTAATGTTAAT 1058

RESULT 8
US-10-650-249-12
; Sequence 12, Application US/10650249
; Publication No. US20040045055A1
; GENERAL INFORMATION:
; APPLICANT: Neff, Michael M
; TITLE OF INVENTION: THE GENE FOR A DOF TRANSCRIPTION FACTOR CAPABLE OF ALTERING
; FILE REFERENCE: WSHU 2064.1
; CURRENT APPLICATION NUMBER: US/10/650,249
; CURRENT FILING DATE: 2003-08-28
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 888
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-650-249-12

Query Match 11.7%; Score 888; DB 7; Length 888;
Best Local Similarity 100.0%; Pred. No. 1.1e-151;
Matches 888; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4232 ATGGTGGAACTGCTCGCATCGCAAAAGTCCCATTCCTGCTGAAGCAGCTCTAAATGGCCCT 4291
Db 1 ATGGTGGAACTGCTCGCATCGCAAAAGTCCCATTCCTGCTGAAGCAGCTCTAAATGGCCCT 60
QY 4292 AGATGTGACTCAACCAATCACTAAGTTCTGTACTTCAATACTATAGCTTACTCAACT 4351
Db 61 AGATGTGACTCAACCAATCACTAAGTTCTGTACTTCAATACTATAGCTTACTCAACT 120
QY 4352 GCCCATTTCTGAAACATGTCGCTATTGGACACGTGGCGGTTCTTGAGGAATGTT 4411
Db 121 GCCCATTTCTGAAACATGTCGCTATTGGACACGTGGCGGTTCTTGAGGAATGTT 180
QY 4412 CTGTTGGAGGAGGCTTTAGGAGGACCAAGAGCAAAATCCAGATCGAAATCTACGGTC 4471
Db 181 CTGTTGGAGGAGGCTTTAGGAGGACCAAGAGCAAAATCCAGATCGAAATCTACGGTC 240
QY 4472 GTGGTCTCGACTGATAACTACTAGTACTTCACTACTTCTTCGCCCCAAGTTACTCA 4531
Db 241 GTGGTCTCGACTGATAACTACTAGTACTTCACTACTTCTTCGCCCCAAGTTACTCA 300
QY 4532 AACCCCTAGCAAGTTTTCATAGCTACGGTCAAATCCCGAGTGTAAATTCCAAATTCGCCATC 4591
Db 301 AACCCCTAGCAAGTTTTCATAGCTACGGTCAAATCCCGAGTGTAAATTCCAAATTCGCCATC 360
QY 4592 TTGCTCTCTCTCAAGCCCTTGAGATTAATTAATTAAGCAACCTGATTAGATTGGT 4651
Db 361 TTGCTCTCTCTCAAGCCCTTGAGATTAATTAATTAAGCAACCTGATTAGATTGGT 420
QY 4652 GGAACCTCAAAATAGCAACATGATAGTGTGATGAGTTCTAGTGGTGGGATCTTGGATGCA 4711
Db 421 GGAACCTCAAAATAGCAACATGATAGTGTGATGAGTTCTAGTGGTGGGATCTTGGATGCA 480
```

```
QY 4712 TGGAATATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTTGATCAACACTACCGGA 4771
Db 481 TGGAATATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTTGATCAACACTACCGGA 540
QY 4772 TTGTTGCAATTTCAAACGGGTTATATCCATTAAGAGTAAGGGAGGTGTTAATCAA 4831
Db 541 TTGTTGCAATTTCAAACGGGTTATATCCATTAAGAGTAAGGGAGGTGTTAATCAA 600
QY 4832 GGTGATTTCTCAACAAGAGTAGTATTTTCCCAATCAGCTAAATGTTTAAAGCCCTTTGATG 4891
Db 601 GGTGATTTCTCAACAAGAGTAGTATTTTCCCAATCAGCTAAATGTTTAAAGCCCTTTGATG 660
QY 4892 GATTTTCTTTCAGGCGGGTTAGCGCCACGCAAAACAAGAAATGTCAAGCGCGAAGAGAAAT 4951
Db 661 GATTTTCTTTCAGGCGGGTTAGCGCCACGCAAAACAAGAAATGTCAAGCGCGAAGAGAAAT 720
QY 4952 GATCAGGATCGGGTAGGGATGGGATGGAATGAATTAATCAAGAAATCTTTTGGGT 5011
Db 721 GATCAGGATCGGGTAGGGATGGGATGGAATGAATTAATCAAGAAATCTTTTGGGT 780
QY 5012 AATATCAACATAAATCAGGCGAGGAGGAGGAAATACATCATCGGGAGGTAAACAGTTCT 5071
Db 781 AATATCAACATAAATCAGGCGAGGAGGAGGAAATACATCATCGGGAGGTAAACAGTTCT 840
QY 5072 TGGACCGGTTTCACTCCCAACAACTCAACAGGCCCATCTCTCATTTCTAA 5119
Db 841 TGGACCGGTTTCACTCCCAACAACTCAACAGGCCCATCTCTCATTTCTAA 888
```

## RESULT 9

```
US-10-225-066A-409
; Sequence 409, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 409
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-409

Query Match 9.1%; Score 692; DB 6; Length 795;
Best Local Similarity 89.5%; Pred. No. 5.7e-116;
Matches 795; Conservative 0; Mismatches 0; Indels 93; Gaps 1;
```

```
QY 4232 ATGTTGGAACGTCTCGGATCGCAAAAGTCCCATTTGCTGAAGCAGCTCTAAATTGCCCT 4291
Db 1 ATGTTGGAACGTCTCGGATCGCAAAAGTCCCATTTGCTGAAGCAGCTCTAAATTGCCCT 60
QY 4292 AGATGTGACTCAACCAATCTAAGTTCTGTTACTTCAATACTATAGCTTACTCAACCT 4351
Db 61 AGATGTGACTCAACCAATCTAAGTTCTGTTACTTCAATACTATAGCTTACTCAACCT 120
QY 4352 CGCCATTTCTGCAAAACATGTCTGCTGCTATTTGGACACGTGGCGGTTCTTGAGGAATGTT 4411
Db 121 CGCCATTTCTGCAAAACATGTCTGCTGCTATTTGGACACGTGGCGGTTCTTGAGGAATGTT 180
QY 4412 CCTGTTGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 4471
Db 181 CCTGTTGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 240
QY 4472 GTGGTCTCGACTGATAAATACTAGTACTTCACTCACTTCTTCGCGCCCAAGTTACTCA 4531
Db 241 GTGGTCTCGACTGATAAATACTAGTACTTCACTCACTTCTTCGCGCCCAAGTTACTCA 300
QY 4532 AACCTAGCAAGTTTTCATAGCTACGGTCAAAATCCCGGAGTTTAAATTCCAAATTCGCCCATC 4591
Db 301 AACCTAGCAAGTTTTCATAGCTACGGTCAAAATCCCGGAGTTTAAATTCCAAATTCGCCCATC 360
QY 4592 TTGCTCTCTCCAAAGCCTTGGAGATTACAATTCGAAGCAACA CTGGATTAGATTGGT 4651
Db 361 TTGCTCTCTCCAAAGCCTTGGAGATTACAATTCGAAGCAACA CTGGATTAGATTGGT 420
QY 4652 GGAATCAATAAGCAACATGATAGTGGTATGAGTTCTAGTGGTGGGATCTTGATGCA 4711
Db 421 GGAATCAATAAGCAACATGATAGTGGTATGAGTTCTAGTGGTGGGATCTTGATGCA 480
QY 4712 TGGAGATACCTCAATCAACAAGCTCAGCAATTCCTTTCTTGATCAACAACATACCGGA 4771
Db 481 TGGAGATACCTCAATCAACAAGCTCAGCAATTCCTTTCTTGATCAACAACATACCGGA 540
QY 4772 TTGTTGCAATCTTCAAAACGCTTATATCCATTTA CTGAAAGGTAAGGAGGTGTTAATCAA 4831
Db 541 TTGTTGCAATCTTCAAAACGCTTATATCCATTTACTAGA----- 578
QY 4832 GGTGATTTCTCAACAGAGTAGTGAATTTTCCATCAGTAAATGTTTAAAGCCCTTGATG 4891
Db 579 ----- 578
QY 4892 GATTTTTCTTCAGGCGGGTTAGCGCCACGCAAAACAAGAAATGTGAAGCGGGAAGAAAT 4951
Db 579 -----AGGCGGGTTAGCGCCACGCAAAACAAGAAATGTGAAGCGGGAAGAAAT 627
QY 4952 GATCAGGATCGGGTAGGGATGGGATGGAGTGAATAAATTATCAAGAAACTTTTGGGT 5011
Db 628 GATCAGGATCGGGTAGGGATGGGATGGAGTGAATAAATTATCAAGAAACTTTTGGGT 687
QY 5012 AATATCAACATAAATCTAGCGGGAACAAGGAAATACATCATGCGGAGGTAACAGTTCT 5071
Db 688 AATATCAACATAAATCTAGCGGGAACAAGGAAATACATCATGCGGAGGTAACAGTTCT 747
QY 5072 TGGACCGGTTTCACTCCAAACATCTCAACAGGCCATCTCTCATTTCAA 5119
Db 748 TGGACCGGTTTCACTCCAAACATCTCAACAGGCCATCTCTCATTTCAA 795
```

## RESULT 10

```
US-10-374-780A-2767
; Sequence 2767, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
```

```
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, I. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Onaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MEI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2767
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1906
US-10-374-780A-2767
```

## Query Match 9.1%; Score 692; DB 7; Length 795;

Best Local Similarity 89.5%; Pred. No. 5.7e-116; Matches 795; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

```
QY 4232 ATGTTGGAACGTCTCGGATCGCAAAAGTCCCATTTGCTGAAGCAGCTCTAAATTGCCCT 4291
Db 1 ATGTTGGAACGTCTCGGATCGCAAAAGTCCCATTTGCTGAAGCAGCTCTAAATTGCCCT 60
QY 4292 AGATGTGACTCAACCAATCTAAGTTCTGTTACTTCAATACTATAGCTTACTCAACCT 4351
Db 61 AGATGTGACTCAACCAATCTAAGTTCTGTTACTTCAATACTATAGCTTACTCAACCT 120
QY 4352 CGCCATTTCTGCAAAACATGTCTGCTGCTATTTGGACACGTGGCGGTTCTTGAGGAATGTT 4411
Db 121 CGCCATTTCTGCAAAACATGTCTGCTGCTATTTGGACACGTGGCGGTTCTTGAGGAATGTT 180
QY 4412 CCTGTTGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 4471
Db 181 CCTGTTGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 240
QY 4472 GTGGTCTCGACTGATAAATACTAGTACTTCACTCACTTCTTCGCGCCCAAGTTACTCA 4531
Db 241 GTGGTCTCGACTGATAAATACTAGTACTTCACTCACTTCTTCGCGCCCAAGTTACTCA 300
QY 4532 AACCTAGCAAGTTTTCATAGCTACGGTCAAAATCCCGGAGTTTAAATTCCAAATTCGCCCATC 4591
Db 301 AACCTAGCAAGTTTTCATAGCTACGGTCAAAATCCCGGAGTTTAAATTCCAAATTCGCCCATC 360
QY 4592 TTGCTCTCTCCAAAGCCTTGGAGATTACAATTCGAAGCAACA CTGGATTAGATTGGT 4651
Db 361 TTGCTCTCTCCAAAGCCTTGGAGATTACAATTCGAAGCAACA CTGGATTAGATTGGT 420
```

4652	Qy	GGAACTCAAAATAGCAACATGATAAGTGGTATGAGTCTTAGTGGTGGGATCTTGGATGCA	4711
421	Db	GGAACTCAAAATAGCAACATGATAAGTGGTATGAGTCTTAGTGGTGGGATCTTGGATGCA	480
4712	Qy	TGGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTCTTCTTGATCAACACACTACCGGA	4771
481	Db	TGGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTCTTCTTGATCAACACTACCGGA	540
4772	Qy	TTGGTGCAAATCTTCAAAACGCGTTATATCCATTACTAGAAAGTAAGGGAGGTGTTAATCAA	4831
541	Db	TTGGTGCAAATCTTCAAAACGCGTTATATCCATTACTAGA-----	578
4832	Qy	GGTGATCTTCAACAGAGAGTAGTGATTAATTCCAATCAGCTAAATGTTTAAGCCCTTGATG	4891
579	Db	-----	578
4892	Qy	GAATTTTCTTCAGCGCGGGTTAGCGCCACGCAACAAGAAATGTGAAGCGGGAAGAGAAT	4951
579	Db	-----AGCGCGGGTTAGCGCCACGCAACAAGAAATGTGAAGCGGGAAGAGAAT	627
4952	Qy	GATCAGGATCGGGTAGGGATGGGGATGGAGTGAATAACTTTATCAAGAACTTTTGGGT	5011
628	Db	GATCAGGATCGGGTAGGGATGGGGATGGAGTGAATAACTTTATCAAGAACTTTTGGGT	687
5012	Qy	AATATCAACATAAACTCAGGCAGGAACGAGGAATACACATCATCGGGAGGTAAACAGTCTCT	5071
688	Db	AATATCAACATAAACTCAGGCAGGAACGAGGAATACACATCATCGGGAGGTAAACAGTCTCT	747
5072	Qy	TGGACCGGTTTCACTCCAAACAACCTCAACAGGCCATCTCTCAATCTCTAA	5119
748	Db	TGGACCGGTTTCACTCCAAACAACCTCAACAGGCCATCTCTCAATCTCTAA	795

## RESULT 11

```

US-10-225-066A-409
; Sequence 409, Application US/10225066A
; Publication No. US20050160493A9
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 409
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-409

```

Query Match	9.1%;	Score 692;	DB 9;	Length 795;
Best Local Similarity	89.5%;	Pred. No. 5.7e-116;		
Matches 795;	Conservative 0;	Mismatches 0;	Indels 93;	Gaps 1;
Qy	4232	ATGGTGGAAAGTGCCTCGGATCGCAAAAGTCCCATTCGCCCTGAAGCAGCTCTAAATTTGCCCT	4291	
Db	1	ATGGTGGAAAGTGCCTCGGATCGCAAAAGTCCCATTCGCCCTGAAGCAGCTCTAAATTTGCCCT	60	
Qy	4292	AGATGTGACTCAACCAATACTAAGTTCTGTACTTCAATAAATATAGCCTTACTCAACCT	4351	
Db	61	AGATGTGACTCAACCAATACTAAGTTCTGTACTTCAATAAATATAGCCTTACTCAACCT	120	
Qy	4352	CGCCATTTCTGCAAAACATGTGCTGCTATTGGACACGTGGCGGTTCCTTGAGGAATGTT	4411	
Db	121	CGCCATTTCTGCAAAACATGTGCTGCTATTGGACACGTGGCGGTTCCTTGAGGAATGTT	180	
Qy	4412	CCGTGTTGGAGGAGGCTTTTAGGAGGAACAAGAGAAACAAATCCAGATCGAAATCTACGGTC	4471	
Db	181	CCGTGTTGGAGGAGGCTTTTAGGAGGAACAAGAGAAACAAATCCAGATCGAAATCTACGGTC	240	
Qy	4472	GTGGTCTCGACTGATTAATACTACTAGTACTTTCATCATCTTACTTCTCGCCCAAGTTACTCA	4531	
Db	241	GTGGTCTCGACTGATTAATACTACTAGTACTTTCATCATCTTACTTCTCGCCCAAGTTACTCA	300	
Qy	4532	AACCTTAGCAAGTTTCATAGCTACGGTCAAATCCCGAGTTTAAATTCGAACCTTGCCTATC	4591	
Db	301	AACCTTAGCAAGTTTCATAGCTACGGTCAAATCCCGAGTTTAAATTCGAACCTTGCCTATC	360	
Qy	4592	TTGGCTCTCTCCAAAGCCCTTGAGAGATTACAAATTCGAAGCAACACTGGATTTAGATTTTGGT	4651	
Db	361	TTGGCTCTCTCCAAAGCCCTTGAGAGATTACAAATTCGAAGCAACACTGGATTTAGATTTTGGT	420	
Qy	4652	GGAACTCAATAAGCAACATGATAAGTGGTATGAGTTCTAGTGGTGGGATCTTGGATGCA	4711	
Db	421	GGAACTCAATAAGCAACATGATAAGTGGTATGAGTTCTAGTGGTGGGATCTTGGATGCA	480	
Qy	4712	TGAGAAATAGCTCCATCAACAAGCTCAGCAATTCCTCTTTTGTATCAACACTACCGGA	4771	
Db	481	TGAGAAATAGCTCCATCAACAAGCTCAGCAATTCCTCTTTTGTATCAACACTACCGGA	540	
Qy	4772	TTGGTGCATCTTTCAACCGGTTATATCCATTAAGAGGTAAAGGAGGTAAATCAA	4831	
Db	541	TTGGTGCATCTTTCAACCGGTTATATCCATTAAGAGGTAAAGGAGGTAAATCAA	578	
Qy	4832	GGTGATTTCCACAGAGAGTAGTGATTAATTCCAATCAGCTAATGTTTAAAGCCCTTGATG	4891	
Db	579	GGTGATTTCCACAGAGAGTAGTGATTAATTCCAATCAGCTAATGTTTAAAGCCCTTGATG	578	
Qy	4892	GATTTTCTTCAGCGGGGTTAGCGCCAACCAAAAGAAATGTGAAGCGGAGAGAGAAT	4951	
Db	579	-----AGCGCGGGTTAGCGCCAACCAAAAGAAATGTGAAGCGGAGAGAGAAT	627	
Qy	4952	GATCAGGATCGGGTAGGGATGGGATGGAATCAATACTTATCAAGAACTTTTGGGT	5011	
Db	628	GATCAGGATCGGGTAGGGATGGGATGGAATGGAATCAATACTTATCAAGAACTTTTGGGT	687	
Qy	5012	AATATCAACATAAATCACTCAGCAGGAAACGAGGAATACATCATCGGGAGGATAACAGTTCT	5071	
Db	688	AATATCAACATAAATCACTCAGCAGGAAACGAGGAATACATCATCGGGAGGATAACAGTTCT	747	
Qy	5072	TGGACCGGTTTCACTCCAAACAATCAACAGGCGATCTCTCATTCTAA	5119	
Db	748	TGGACCGGTTTCACTCCAAACAATCAACAGGCGATCTCTCATTCTAA	795	

## RESULT 12

US-10-650-249-15  
; Sequence 15, Application US/10650249  
; Publication No. US20040045055A1  
; GENERAL INFORMATION:  
; APPLICANT: Neff, Michael M  
; TITLE OF INVENTION: THE GENE FOR A  
; TITLE OF INVENTION: THE SIZE AND S

FILE REFERENCE: WSHU 2064.1  
CURRENT APPLICATION NUMBER: US/10/650,249  
CURRENT FILING DATE: 2003-08-28  
PRIOR APPLICATION NUMBER: US 60/406,657  
PRIOR FILING DATE: 2002-08-02  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 15  
LENGTH: 684  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-10-650-249-15

Query Match 9.0%; Score 684; DB 7; Length 684;  
Best Local Similarity 100.0%; Pred. No. 1.5e-114;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	4436	AACAGAGAGCAATCCAGATCGAAATCTACGGTCGGTCTCGACTGATAATACTACT	4495
DB	1	AACAAGAGCAATCCAGATCGAAATCTACGGTCGGTCTCGACTGATAATACTACT	60
QY	4496	AGTACTTTCATCACTTACTTCTCGCCCAAGTTACTCAAAACCCCTAGCAAGTTTCATAGCTAC	4555
DB	61	AGTACTTTCATCACTTACTTCTCGCCCAAGTTACTCAAAACCCCTAGCAAGTTTCATAGCTAC	120
QY	4556	GGTCAAAATCCCGAGTTTAATTCAACTTGCCATCTTGCCTCTCTCCAAAGCCTTGA	4615
DB	121	GGTCAAAATCCCGAGTTTAATTCAACTTGCCATCTTGCCTCTCTCCAAAGCCTTGA	180
QY	4616	GATTACAATCAAGCAACACTGGATTAGATTTTGGTGGAACTCAAAATAGCAACATGATA	4675
DB	181	GATTACAATCAAGCAACACTGGATTAGATTTTGGTGGAACTCAAAATAGCAACATGATA	240
QY	4676	AGTGTATGAGTCTTAGTGGTGGATCTTGGATGATGGAGAAATACCTCCATCAACAA	4735
DB	241	AGTGTATGAGTCTTAGTGGTGGATCTTGGATGATGGAGAAATACCTCCATCAACAA	300
QY	4736	GCTCAGCAATTCCTTTCTTGATCAACACTACCGGATTTGGTGAATCTTCAAAACCGTTA	4795
DB	301	GCTCAGCAATTCCTTTCTTGATCAACACTACCGGATTTGGTGAATCTTCAAAACCGTTA	360
QY	4796	TATCAATCTAGAGGTAAGGAGTGTAAATCAAGGTGATCTCAACAGAGTAGT	4855
DB	361	TATCAATCTAGAGGTAAGGAGTGTAAATCAAGGTGATCTCAACAGAGTAGT	420
QY	4856	GATTATTCCAATCAGCTAATGTTTAAGCCCTTGATGGATTTTCTCAGGCGGGTTAGC	4915
DB	421	GATTATTCCAATCAGCTAATGTTTAAGCCCTTGATGGATTTTCTCAGGCGGGTTAGC	480
QY	4916	GCCAGCAAAACAAAGAAATGTGAAGCGGAAGAGAAATGATCAGGATCGGGTAGGGATGG	4975
DB	481	GCCAGCAAAACAAAGAAATGTGAAGCGGAAGAGAAATGATCAGGATCGGGTAGGGATGG	540
QY	4976	GATGAGTGAATAACTTATCAAGAAACCTTTGGGTAAATATCAACATAACTCAGGCAGG	5035
DB	541	GATGAGTGAATAACTTATCAAGAAACCTTTGGGTAAATATCAACATAACTCAGGCAGG	600
QY	5036	AACGAGGAATACACATATGCGGAGGTAAACAGTTCTTGACCGGTTTCACCTCCAAACAC	5095
DB	601	AACGAGGAATACACATATGCGGAGGTAAACAGTTCTTGACCGGTTTCACCTCCAAACAC	660
QY	5096	TCAACAGGCCATCTCTCAATCTTAA	5119
DB	661	TCAACAGGCCATCTCTCAATCTTAA	684

## RESULT 13

US-10-161-403-90  
Sequence 90, Application US/10161403  
Publication No. US20030119104A1  
GENERAL INFORMATION:  
APPLICANT: Perkins, Edward  
APPLICANT: Perez, Carl

APPLICANT: Lindenbaum, Michael  
APPLICANT: Greene, Amy  
APPLICANT: Leung, Josephine  
APPLICANT: Fleming, Elena  
APPLICANT: Stewart, Sandra  
APPLICANT: Shellard, Joan  
TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
FILE REFERENCE: 24601-420  
CURRENT APPLICATION NUMBER: US/10/161,403  
CURRENT FILING DATE: 2002-05-30  
PRIOR APPLICATION NUMBER: 60/294,758  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/366,891  
PRIOR FILING DATE: 2002-03-21  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 90  
LENGTH: 8428  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: pCambia3300 Plasmid  
US-10-161-403-90

Query Match 8.1%; Score 617.2; DB 6; Length 8428;

Best Local Similarity	93.6%;	Pred. No. 7.3e-33;	Mismatches	0;	Mismatches	33;	Indels	12;	Gaps	1;
Matches	659;	Conservative	0;							
QY	6868	AGATATACATCAATCCACTTGTCTTGAAGACGTGGTTGGAAACGCTCTCTTTTCCACGA	6927							
DB	7146	AGATATACATCAATCCACTTGTCTTGAAGACGTGGTTGGAAACGCTCTCTTTTCCACGA	7205							
QY	6928	TGTTCTCTCGTGGTGGGTCCATCTTTGGGACCACTCTCGGTAGAGCATCTTGAAACGA	6987							
DB	7206	TGTTCTCTCGTGGTGGGTCCATCTTTGGGACCACTCTCGGTAGAGCATCTTGAAACGA	7265							
QY	6988	TAGCTCTCTCTTATCGCAATGATGGCAATTTGTAGGAGCCATCTCTCTTCTACTGTCC	7047							
DB	7266	TAGCTCTCTCTTATCGCAATGATGGCAATTTGTAGGAGCCATCTCTCTTCTACTGTCC	7325							
QY	7048	TTTCGATGAAGTGACAGATAGCTGGGCAATCGAATCGAATCGGAGGTTTCCCGATATTACCC	7107							
DB	7326	TTTCGATGAAGTGACAGATAGCTGGGCAATCGAATCGAATCGGAGGTTTCCCGATATTACCC	7385							
QY	7108	TTTGTGAAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATCTTGG	7167							
DB	7386	TTTGTGAAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATCTTGG	7445							
QY	7168	AGTAGACGAGAGTGTCTGTCTCCACCATGTTGGGGATCTAGATATCAATCAATCCACTT	7227							
DB	7446	AGTAGACGAGAGTGTCTGTCTCCACCATGTTGGGGATCTAGATATCAATCAATCCACTT	7493							
QY	7228	GCTTTGAAGACGTGGTTGGAACGCTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTTC	7287							
DB	7494	GCTTTGAAGACGTGGTTGGAACGCTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTTC	7553							
QY	7288	CATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCCTTCTTTTATCGCAAT	7347							
DB	7554	CATCTTTGGGACCACTGTCTGGTAGAGGCATCTTGAACGATAGCCCTTCTTTTATCGCAAT	7613							
QY	7348	GATGCAATTTGTAGAGGCATCTTCTCTTCTACTGTCTCTTTCGATGAGTACAGATAG	7407							
DB	7614	GATGCAATTTGTAGAGGCATCTTCTCTTCTACTGTCTCTTTCGATGAGTACAGATAG	7673							
QY	7408	CTGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCTTTTGTGAAAAAGTCTCAATAG	7467							
DB	7674	CTGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCTTTTGTGAAAAAGTCTCAATAG	7733							
QY	7468	CCCTCTGGTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGAGTCTCGTCT	7527							
DB	7734	CCCTCTGGTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGAGTCTCGTCT	7793							
QY	7528	CCACCATGTTGGGATTCACCTAGTCTTAGAGCGGCCGCCACCGC	7571							







Qy	7288	CATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAAACGATAGCCCTTTCCTTTATCGCAAT	7347
Db	7554	CATCTTTGGGACCACTGTCGGCAGAGGCATCTTGAAACGATAGCCCTTTCCTTTATCGCAAT	7613
Qy	7348	GATGGCAATTTGTAGAACCATCTTCCTTTTCTACTGTCTCTTCGATGAAGTGACAGATAG	7407
Db	7614	GATGGCAATTTGTAGGTGCCCTTCCTTTTCTACTGTCTCTTCGATGAAGTGACAGATAG	7673
Qy	7408	CTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCCCTTTGTTGAAAAGTCTCAATAG	7467
Db	7674	CTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCCCTTTGTTGAAAAGTCTCAATAG	7733
Qy	7468	CCCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGAGTGTGCTGCT	7527
Db	7734	CCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGAGTGTGCTGCT	7793
Qy	7528	CCACCATGTTGGGATCCACTAGTTCTAGAGCGGCCGCCGC	7571
Db	7794	CCACCATGTTGGCAAGCTGCTCTAGCCAATACGCAACCGCCTC	7837

Search completed: December 31, 2005, 06:23:45  
Job time : 3577 secs

**BEST AVAILABLE COPY**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2005, 16:31:40 ; Search time 840 Seconds  
(without alignments)  
16040.378 Million cell updates/sec

Title: US-10-650-249-1  
Perfect score: 7580  
Sequence: 1 agctctataattcaagaga.....gccgccacccgggtgagact 7580

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/PCBUS COMB.seq:\*  
7: /cgn2\_6/prodata/1/ina/PP COMB.seq:\*  
8: /cgn2\_6/prodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	621	8.2	5033	2	US-08-038-768A-1
C 2	617.2	8.1	12614	3	US-09-577-424-1
C 3	563	7.4	2728	3	US-08-836-402B-7
C 4	554.2	7.3	1138	3	US-09-011-151-8
C 5	554.2	7.3	1138	3	US-09-011-151-9
C 6	462.2	6.1	907	3	US-09-623-551-17
C 7	461	6.1	5767	3	US-09-810-861B-3
C 8	461	6.1	14446	3	US-09-810-861B-4
C 9	379.8	5.0	1800	3	US-09-737-698B-29
C 10	379.8	5.0	1800	3	US-09-737-626A-29
C 11	379.8	5.0	1800	3	US-10-427-169-29
C 12	379.8	5.0	1800	3	US-10-427-180-29
C 13	379.2	5.0	1742	3	US-09-737-698B-30
C 14	379.2	5.0	1742	3	US-09-737-626A-30
C 15	379.2	5.0	1742	3	US-10-427-169-30
C 16	379.2	5.0	1742	3	US-10-427-180-30
C 17	377.6	5.0	661	3	US-09-027-998A-33
C 18	377.6	5.0	661	3	US-09-943-692-33
C 19	377.6	5.0	2107	3	US-09-441-340-29
C 20	377.6	5.0	2122	3	US-09-441-340-25
C 21	377.6	5.0	2436	3	US-09-441-340-31
C 22	377.6	5.0	3469	3	US-09-377-466B-23
C 23	377.6	5.0	3469	3	US-10-232-665-23
C 24	377.6	5.0	3754	3	US-09-377-466B-15

C 25	377.6	5.0	3754	3	US-10-232-665-15
C 26	377.6	5.0	8012	3	US-09-182-117-1
C 27	377.6	5.0	8012	3	US-09-434-039A-1
C 28	377.6	5.0	8349	3	US-09-186-002-16
C 29	377.6	5.0	8418	3	US-09-182-117-5
C 30	377.6	5.0	8418	3	US-09-434-039A-5
C 31	377.6	5.0	8798	3	US-09-182-117-4
C 32	377.6	5.0	8798	3	US-09-434-039A-4
C 33	377.6	5.0	10249	3	US-09-186-002-14
C 34	377.6	5.0	10339	3	US-09-186-002-13
C 35	374.6	4.9	1030	2	US-07-936-163-46
C 36	374.6	4.9	1030	3	US-08-729-601A-43
C 37	374.6	4.9	1196	3	US-08-729-601A-46
C 38	374.6	4.9	5897	3	US-09-097-319A-26
C 39	374.6	4.9	5897	3	US-09-643-971-26
C 40	374.6	4.9	9335	3	US-09-097-319A-19
C 41	374.6	4.9	9335	3	US-09-643-971-19
C 42	374.6	4.9	10160	3	US-09-097-319A-8
C 43	374.6	4.9	10160	3	US-09-643-971-8
C 44	374.6	4.9	11784	3	US-09-097-319A-9
C 45	374.6	4.9	11784	3	US-09-643-971-9

ALIGNMENTS

RESULT 1  
US-08-038-768A-1/c  
; Sequence 1. Application US/08038768A  
; Patent No. 5569828  
; GENERAL INFORMATION:  
; APPLICANT: McMullen, Michael D.; Roth, Bradley A.; Townsend,  
; APPLICANT: Rod  
; TITLE OF INVENTION: MAIZE CHLOROTIC DWARF VIRUS RESISTANCE  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pioneer Hi-Bred International, Inc.  
; STREET: 700 Capital Square, 400 Locust  
; STREET: Street  
; CITY: Des Moines  
; STATE: Iowa  
; COUNTRY: United States  
; ZIP: 50309  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS, Microsoft Windows  
; SOFTWARE: Microsoft Windows No. 5569828epad  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/038,768A  
; FILING DATE: 19930324  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Roth, Michael J.  
; REGISTRATION NUMBER: 29,342  
; REFERENCE/DOCKET NUMBER: 0235 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (515) 245-3594  
; TELEFAX: (515) 245-3634  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5033 base pairs  
; TYPE: nucleotide  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: synthetic DNA  
; DESCRIPTION: transformation plasmid pPH11406  
; HYPOTHETICAL: NO  
US-08-038-768A-1

```
Query Match      8.2%; Score 621; DB 2; Length 5033;
Best Local Similarity 95.1%; Pred. No. 2.8e-111;
Matches 653; Conservative 0; Mismatches 30; Indels 4; Gaps 1;

QY 6868 AGATATACATCAATCCACTTCTTTGAAGACGTGGTGGAAACGTCTCTTTTCCACGA 6927
Db 1103 AGATATACATCAATCCACTTCTTTGAAGACGTGGTGGAAACGTCTCTTTTCCACGA 1044
QY 6928 TGTTCCTCGTGGTGGGGTCCATCTTTGGGACCACTGTCGTAGAGGCAATCTTGAACGA 6987
Db 1043 TGTTCCTCGTGGTGGGGTCCATCTTTGGGACCACTGTCGTAGAGGCAATCTTGAACGA 984
QY 6988 TAGCCCTTCTTTATCGCAATGATGCAATTTGTAGAAGCACTCTCTCTTCTACTGTCC 7047
Db 983 TAGCCCTTCTTTATCGCAATGATGCAATTTGTAGAAGCACTCTCTCTTCTACTGTCC 924
QY 7048 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCC 7107
Db 923 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCC 864
QY 7108 TTTTGTGAAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTGATATCTTGG 7167
Db 863 TTTTGTGAAAAGTCTCAATAGCCCTTGTGCTCTCTGAGACTGTATCTTGATATCTTGG 804
QY 7168 AGTAGACGAGAGTGTGCTGCCACCACTGTGGGATCT---AGATATACATCAATCC 7223
Db 803 AGTAGACGAGAGTGTGCTGCCACCACTGTGACGATTTCTAGAGCATCAATCC 744
QY 7224 ACTTGTGTTGAAGACGTGGTGGAAACGTCTCTTTTCCACGATGTTCTCGTGGTGGG 7283
Db 743 ACTTGTGTTGAAGACGTGGTGGAAACGTCTCTTTTCCACGATGTTCTCGTGGTGGG 684
QY 7284 GGTCCATCTTTGGGACCACTGTCGTAGAGGCAATCTTGAACGATAGCCCTTCTTTATCG 7343
Db 683 GGTCCATCTTTGGGACCACTGTCGCAGAGGCAATCTTGAACGATAGCCCTTCTTTATCG 624
QY 7344 CAATGATGCAATTTGTAGAAGCACTCTCTCTTTCTACTGTCTCTTGCATGAAGTGACAG 7403
Db 623 CAATGATGCAATTTGTAGTGCCACCTTCTTTTCTACTGTCTCTTGCATGAAGTGACAG 564
QY 7404 ATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGAAAAGTCTCA 7463
Db 563 ATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGAAAAGTCTCA 504
QY 7464 ATAGCCCTCGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGCG 7523
Db 503 ATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGCG 444
QY 7524 TGTCCACCATGTTGGGATCCACTAG 7550
Db 443 TGTCCACCATGTTGACGATTCTAG 417
```

```
RESULT 2
US-09-577-424-1/c
; Sequence 1, Application US/09577424
; Patent No. 6525245
; GENERAL INFORMATION:
; APPLICANT: Rhoads, David M
; TITLE OF INVENTION: METHOD FOR IDENTIFYING COMPONENTS INVOLVED IN SIGNAL
; TITLE OF INVENTION: TRANSDUCTION PATHWAYS IN HIGHER PLANTS
; FILE REFERENCE: UNL2990
; CURRENT APPLICATION NUMBER: US/09/577,424
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 12614
; TYPE: DNA
; ORGANISM: Bacterial Plasmid DNA - Plasmid pZP-ALE
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1665)..(3317)
```

```
US-09-577-424-1
Query Match      8.1%; Score 617.2; DB 3; Length 12614;
Best Local Similarity 93.6%; Pred. No. 1.9e-110;
Matches 659; Conservative 0; Mismatches 33; Indels 12; Gaps 1;

QY 6868 AGATATACATCAATCCACTTCTTTGAAGACGTGGTGGAAACGTCTCTTTTCCACGA 6927
Db 4915 AGATATACATCAATCCACTTCTTTGAAGACGTGGTGGAAACGTCTCTTTTCCACGA 4856
QY 6928 TGTTCCTCGTGGTGGGGTCCATCTTTGGGACCACTGTCGTAGAGGCAATCTTGAACGA 6987
Db 4855 TGTTCCTCGTGGTGGGGTCCATCTTTGGGACCACTGTCGTAGAGGCAATCTTGAACGA 4796
QY 6988 TAGCCCTTCTTTATCGCAATGATGCAATTTGTAGAAGCACTCTCTTCTACTGTCC 7047
Db 4795 TAGCCCTTCTTTATCGCAATGATGCAATTTGTAGTGCCACTTCTCTTCTACTGTCC 4736
QY 7048 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCC 7107
Db 4735 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCC 4676
QY 7108 TTTTGTGAAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTGATATCTTGG 7167
Db 4675 TTTTGTGAAAAGTCTCAATAGCCCTTGTGCTCTCTGAGACTGTATCTTGATATCTTGG 4616
QY 7168 AGTAGACGAGAGTGTGCTGCCACCACTGTGGGATCTTAGATATCAATCAATCACTT 7227
Db 4615 AGTAGACGAGAGTGTGCTGCCACCACTGT-----TATCAATCAATCACTT 4568
QY 7228 GCTTTGAAGACGTGGTGGAAACGTCTCTTTTCCACGATGTTCTCGTGGTGGGGTCT 7287
Db 4567 GCTTTGAAGACGTGGTGGAAACGTCTCTTTTCCACGATGTTCTCGTGGTGGGGTCT 4508
QY 7288 CATCTTTGGGACCACTGTCGTAGAGGCAATCTTGAACGATAGCCCTTCTTTATGCAAT 7347
Db 4507 CATCTTTGGGACCACTGTCGCAGAGGCAATCTTGAACGATAGCCCTTCTTTATGCAAT 4448
QY 7348 GATGCAATTTGTAGAAGCACTCTCTTTCTACTGTCTCTTTCGATGAAGTACAGATAG 7407
Db 4447 GATGCAATTTGTAGTGCCACTTCTCTTTCTACTGTCTCTTTGATGAAGTACAGATAG 4388
QY 7408 CTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGAAAAGTCTCAATAG 7467
Db 4387 CTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCTTTGTTGAAAAGTCTCAATAG 4328
QY 7468 CCTCTGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGTCGTCT 7527
Db 4327 CCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGTCGTCT 4268
QY 7528 CCACCATGTTGGGATCCACTAGTTCTAGAGCGCGCCACCGC 7571
Db 4267 CCACCATGTTGGCAAGCTGCTCTAGCCCAATACGCAACCGCCTC 4224
```

```
RESULT 3
US-08-836-402B-7/c
; Sequence 7, Application US/08836402B
; Patent No. 6063988
; GENERAL INFORMATION:
; APPLICANT: Rudiger Hain, Regina Fischer
; TITLE OF INVENTION: DNA SEQUENCE AND ITS USE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
; COMPUTER: Gateway 2000
```

```
; OPERATING SYSTEM: DOS
; SOFTWARE: Microsoft Windows 98
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,402B
; FILING DATE: 02-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 44 40 200.7 (Germany)
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: Bayer 9840-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2728 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; US-08-836-402B-7

Query Match 7.4%; Score 563; DB 3; Length 2728;
Best Local Similarity 97.7%; Pred. No. 4.6e-100;
Matches 607; Conservative 0; Mismatches 0; Indels 14; Gaps 3;

QY 6190 AGATATCACATCAATCCACTGCTTTGAAGACGTCGTTGGAACGTCCTCTTTTCCACGA 6249
DB 627 AGATATCACATCAATCCACTGCTTTGAAGACGTCGTTGGAACGTCCTCTTTTCCACGA 568
QY 6250 TCTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGAGGCA-TCTTGAACG 5308
DB 567 TCTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGAGGCAATCTTGAACG 508
QY 6309 ATAGCCTTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTCTACTGTC 6368
DB 507 ATAGCCTTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTCTACTGTC 448
QY 6369 CTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACC 6428
DB 447 CTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACC 388
QY 6429 CTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTTG 6488
DB 387 CTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTTG 328
QY 6489 GAGTAGACGAGAGTGTGCTGCCACCATGTTGGGATCTAGATATCACATCAATCCACT 6548
DB 327 GAGTAGACGAGAGTGTGCTGCCACCATGT-----TATCATCATCAATCCACT 280
QY 6549 TGCCTTTGAAGACGTCGTTGGAACGTCCTCTTTTCCACGATGTTCTCGTGGGTGGGGT 6608
DB 279 TGCCTTTGAAGACGTCGTTGGAACGTCCTCTTTTCCACGATGTTCTCGTGGGTGGGGT 220
QY 6609 CCATCTTTGGGACCACTGTCGAGTACAGGCA-TCTTGAACGATAGCTTCTTTTATCGCA 6667
DB 219 CCATCTTTGGGACCACTGTCGAGTACAGGCAATCTTGAACGATAGCTTCTTTTATCGCA 160
QY 6668 ATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCCTTTTCGATGAAGTACAGAT 6727
DB 159 ATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCCTTTTCGATGAAGTACAGAT 100
QY 6728 AGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCTTTGTGAAAAGTCTCAAT 6787
DB 99 AGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCTTTGTGAAAAGTCTCAAT 40
QY 6788 AGCCCTCTGGTCTCTGAGAC 6808
|||||
```

```
DB 39 AGCCCTCTGGTCTTCTGAGAC 19
RESULT 4
US-09-011-151-8/c
; Sequence 8, Application US/09011151
; Patent No. 6380463
; GENERAL INFORMATION:
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: DNA Constructs
; FILE REFERENCE: PPD 50059/UST
; CURRENT APPLICATION NUMBER: US/09/011,151
; CURRENT FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: PCT/GB96/01893
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: GB 9515941.4
; PRIOR FILING DATE: 1995-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 8
; LENGTH: 1138
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Plasmid pMJB1
; US-09-011-151-8

Query Match 7.3%; Score 554.2; DB 3; Length 1138;
Best Local Similarity 92.3%; Pred. No. 1.9e-98;
Matches 621; Conservative 0; Mismatches 38; Indels 14; Gaps 3;

QY 6190 AGATATCACATCAATCCACTGCTTTGAAGACGTCGTTGGAACGTCCTCTTTTCCACG 6248
DB 681 AGATATCACATCAATCCACTGCTTTGAAGACGTCGTTGGAACGTCCTCTTTTCCACG 622
QY 6249 ATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGAGGCAATCTTGAACG 6308
DB 621 ATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGAGGCAATCTTGAACG 562
QY 6309 ATAGCCTTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTCTACTGTC 6368
DB 561 ATAGCCTTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTCTACTGTC 502
QY 6369 CTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACC 6428
DB 501 TTTCAATAAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACC 442
QY 6429 CTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTTG 6488
DB 441 CTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTTG 382
QY 6489 GAGTAGACGAGAGTGTGCTGCCACCATGTTGGGATCTAGATATCACATCAATCCACT 6548
DB 381 GAGTAGACGAGAGTGTGCTGCCACCATGT-----TATCATCATCAATCCACT 334
QY 6549 TGCCTTTGAAGACGTCGTTGGAACGTCCTCTTTTCCACGATGTTCTCGTGGGTGGGGT 6607
DB 333 TGCCTTTGAAGACGTCGTTGGAACGTCCTCTTTTCCACGATGTTCTCGTGGGTGGGGT 274
QY 6608 TCCATCTTTGGGACCACTGTCGAGTACAGGCAATCTTGAACGATAGCTTCTTTTATCGCA 6667
DB 273 TCCATCTTTGGGACCACTGTCGAGTACAGGCAATCTTGAACGATAGCTTCTTTTATCGCA 214
QY 6668 ATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCCTTTTCGATGAAGTACAGAT 6727
DB 213 ATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCCTTTTCGATGAAGTACAGAT 154
QY 6728 AGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCTTTGTGAAAAGTCTCAAT 6787
DB 153 AGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCTTTGTGAAAAGTCTCAAT 94
QY 6788 AGCCCTCTGGTCTCTGAGACTGTATCTTTGATATTCTTGGAGTACGAGAGTCTGCTG 6847
|||||
```

```
Db 93 TGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATATTTTGGAGTAGACAAGTGTGCTG 34
Qy 6848 CTCACCATGTTG 6860
Db 33 CTCACCATGTTG 21

RESULT 5
US-09-011-151-9
; Sequence 9, Application US/09011151
; Patent No. 6380463
; GENERAL INFORMATION:
; APPLICANT: Jenson, Ian
; TITLE OF INVENTION: DNA Constructs
; FILE REFERENCE: PPD 50059/UST
; CURRENT APPLICATION NUMBER: US/09/011,151
; CURRENT FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: PCT/GB96/01883
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: GB 9515941.4
; PRIOR FILING DATE: 1995-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1138
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Plasmid pMJBI
US-09-011-151-9

Query Match 7.3%; Score 554.2; DB 3; Length 1138;
Best Local Similarity 92.3%; Pred. No. 1.9e-98;
Matches 621; Conservative 0; Mismatches 38; Indels 14; Gaps 3;

Qy 6190 AGATATCATCAATCAATCCACTTGTCTTTGAGAGCTGGTTGGACGCTTTC-TTTTCCACG 6248
Db 458 AGATATCATCAATCAATCCACTTGTCTTTGAGAGCTGGTTGGACGCTTTC-TTTTCCACG 517

Qy 6249 ATGTTCTCTGCTGGGTGGGGTCCATCTTTGGACCACTGTCCGTTAGAGGCACTTTGAACG 6308
Db 518 ATGCTCTCTGCTGGGTGGGGTCCATCTTTGGACCACTGTCCGTTAGAGGCACTTTCAAG 577

Qy 6309 ATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTTTCTACTGTC 6368
Db 578 ATGGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTTTCTCACTATC 637

Qy 6369 CTTTCGATCAAGTGACAGTAGCTGGCAATCGAATCCGAGAGGTTTCCCGATATTACC 6428
Db 638 TTCACAATAAAGTGACAGTAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACC 697

Qy 6429 CTTTGTGTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTCTTG 6488
Db 698 CTTTGTGTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTCTG 757

Qy 6489 GAGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGGAATAGATATCAATCAATCCACT 6548
Db 758 GAGTAGACAAAGTGTGCTGCTCCACCATGT-----TATCAATCAATCCACT 805

Qy 6549 TGCCTTTGAAGAGCTGGTTGGAAGCTTTC-TTTTCCAGATGTTCCCTGCTGGGTGGGG 6607
Db 806 TGCCTTTGAAGAGCTGGTTGGAAGCTTTC-TTTTCCAGATGTTCCCTGCTGGGTGGGG 865

Qy 6608 TCCATCTTTGGGACCACTGTCCGTTAGAGGCACTTTGGAAGATAGCCCTTTCTTTATCGCA 6667
Db 866 TCACTCTTTGGGACCACTGTCCGTTAGAGGCACTTTGGAAGATAGCCCTTTCTTTATCGCA 925

Qy 6668 ATGATGGCAATTTGTAGAGCCATCTTCTCTTCTACTGCTTTCGATCAAGTGACAGAT 6727
Db 926 ATGATGGCAATTTGTAGAGCCATCTTCTCTTCTACTGCTTTCGATCAAGTGACAGAT 985

Qy 6728 AGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCCTTTGTTGAAAAGTCTCAAT 6787
Db 926 AGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCCTTTGTTGAAAAGTCTCAAT 6787
```

```
Db 986 AGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCCTTTGTTGAAAAGTCTCAAT 1045
Qy 6788 AGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATCTTTGGAGTAGACAAGTGTGCTG 6847
Db 1046 TGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATATTTTGGAGTAGACAAGTGTGCTG 1105
Qy 6848 CTCACCATGTTG 6860
Db 1106 CTCACCATGTTG 1118

RESULT 6
US-09-623-551-17/c
; Sequence 17, Application US/09623551
; Patent No. 6774281
; GENERAL INFORMATION:
; APPLICANT: Stuiver, Maarten
; APPLICANT: Custers, Jerome
; APPLICANT: Simons, Lambertus
; TITLE OF INVENTION: Method For The Induction Of Pathogen Resistance In
; TITLE OF INVENTION: Plants
; FILE REFERENCE: MOG57706/UST
; CURRENT APPLICATION NUMBER: US/09/623,551
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: EP 98104076.9
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: PCT/EP99/01672
; PRIOR FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 907
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Combination of
; OTHER INFORMATION: 35S Promoter with the Glucanase II leader
US-09-623-551-17

Query Match 6.1%; Score 462.2; DB 3; Length 907;
Best Local Similarity 80.4%; Pred. No. 1.4e-80;
Matches 602; Conservative 0; Mismatches 43; Indels 104; Gaps 1;

Qy 6190 AGATATCAATCAATCCACTTGTCTTTGAGAGCTGGTTGGAACTGCTTTTTCACGA 6249
Db 766 AGATATCAATCAATCCACTTGTCTTTGAGAGCTGGTTGGAACTGCTTTTTCACGA 707

Qy 6250 TGTTCCTCTGCTGGGTGGGGTCCATCTTTGGGACCACTGTCCGTTAGAGGCACTTTGAACGA 6309
Db 706 TGTTCCTCTGCTGGGTGGGGTCCATCTTTGGGACCACTGTCCGTTAGAGGCACTTTGAACGA 647

Qy 6310 TAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAGGCACTTCTTCTTCTACTGTC 6369
Db 646 TGGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAGGCACTTCTTCTTCTACTGTC 587

Qy 6370 TTTTCATGAGTAGACAGTAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACC 6429
Db 586 TCACAATAAAGTGACAGTAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACC 527

Qy 6430 TTTTGTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTCTTG 6489
Db 526 TTTTGTGAAAAGTCTCAATAGCCCTTGGTCTTCTGAGACTGTATCTTTGATATTCTTG 467

Qy 6490 AGTAGACGAGAGTGTGCTGCTCCACATGTTGGGA----- 6525
Db 466 AGTAGACAAAGTGTGCTGCTCCACATGTTGACGAAGATTTTCTTCTTGTCAATTGAGTC 407

Qy 6526 ----- 6525
Db 406 GTAAGAGACTCTGTATGAACCTGTTCCGCCAGTCTTTACGGCGAGTCTGTAGGCTCTCTA 347

Qy 6526 -----TCTAGATATCAATCAATCCACTTCTTCTTGTGAAAGCTGGT 6565
Db 6526 -----TCTAGATATCAATCAATCCACTTCTTCTTGTGAAAGCTGGT 6565
```

Db 346 TTGTGAATCTTTGACTCCATGCGGAATTATCATCATCAATCCACTTGTCTTTGAAGACGTGT 287  
QY 6566 TGGAAACGTCTCTTTTTCACAGATGTTCTCTGGGTGGGGTCCATCTTTTGGGACCACT 6625  
Db 286 TGGAAACGTCTCTTTTTCACAGATGTTCTCTGGGTGGGGTCCATCTTTTGGGACCACT 227  
QY 6626 GTCCGTAGAGGATCTTGAACAGATAGCCCTTCTCTTATCGCAATGATGGCAATTTAGAA 6685  
Db 226 GTCCGAGAGGATCTTCAACAGATGCGCTTCTCTTATCGCAATGATGGCAATTTAGAA 167  
QY 6686 GCATCTCTCTTTCTACTGTCTCTTGAATGAAGTGACAGATAGCTGGGCAATGGAATCC 6745  
Db 166 GCACCTCTCTTTTTCACATCTTACAAATAAGTGACAGATAGCTGGGCAATGGAATCC 107  
QY 6746 GAGGAGTTTCCGATATTACCTTTTGTGAAAGTCTCAATAGCCCTCTGTCTCTGA 6805  
Db 106 GAGGAGTTTCCGATATTACCTTTTGTGAAAGTCTCAATAGCCCTTTGTCTCTGA 47  
QY 6806 GACTGTATCTTTGATATCTTTGGAGTAGA 6834  
Db 46 GACTGTATCTTTGATATCTTTGGAGTAGA 18  
RESULT 7  
US-09-810-861B-3/c  
; Sequence 3, Application US/09810861B  
; Patent No. 6770799  
; GENERAL INFORMATION:  
; APPLICANT: Mor, Tsafir S.  
; APPLICANT: Soreq, Hermona  
; APPLICANT: Arntzen, Charles J.  
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN  
; TITLE OF INVENTION: TRANSGENIC PLANTS  
; FILE REFERENCE: BTI-45  
; CURRENT APPLICATION NUMBER: US/09/810,861B  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/190,440  
; PRIOR FILING DATE: 2000-03-17  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 3  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: plasmid vector  
; OTHER INFORMATION: pTM034.  
US-09-810-861B-3  
Query Match 6.1%; Score 461; DB 3; Length 5767;  
Best Local Similarity 86.5%; Pred. No. 3.9e-80;  
Matches 582; Conservative 0; Mismatches 0; Indels 91; Gaps 3;  
QY 6190 AGATATCATCATCAATCCACTTGTGAAAGACGTGTGGAACGTCCTTTTTCACGA 6249  
Db 603 AGATATCATCATCAATCCACTTGTGAAAGACGTGTGGAACGTCCTTTTTCACGA 544  
QY 6250 TGTTCTCTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCA-TCTTGAACG 6308  
Db 543 TGTTCTCTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCAATCTTGAACG 484  
QY 6309 ATAGCTTTCTTTATCGCAATGATGGCAATTTGTGAAGCCATCTTCTTTTCTACTGTC 6368  
Db 483 ATAGCTTTCTTTATCGCAATGATGGCAATTTGTGAAGCCATCTTCTTTTCTACTGTC 424  
QY 6369 CTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACC 6428  
Db 423 CTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACC 364  
QY 6429 CTTTGTGAAAGTCTCAATAGCCCTCTGTCTTCTGAGACTGTATCTTTGATATCTTGTG 6488  
Db 363 CTTTGTGAAAGT----- 350

QY 6489 GAGTAGACGAGAGTGTGTGCTCCACCATGTTGGGATCTAGATATCATCATCAATCCACT 6548  
Db 349 -----TATCATCATCAATCCACT 333  
QY 6549 TGCTTTGAAGACGTGTGGAACGTCCTTTTTCACAGATGTTCTCTGGGTGGGGT 6608  
Db 332 TGCTTTGAAGACGTGTGGAACGTCCTTTTTCACAGATGTTCTCTGGGTGGGGT 273  
QY 6609 CCATCTTTGGGACCACTGTCGGTAGAGGCA-TCTTGAAGCATAGAGCTTTCTTTATCGCA 6667  
Db 272 CCATCTTTGGGACCACTGTCGGTAGAGGCAATCTTTGAAGCATAGAGCTTTCTTTATCGCA 213  
QY 6668 ATGATGGCAATTTGTGAAGCCATCTCTTTTCTACTGTCCTTTTCGATGAAGTGACAGAT 6727  
Db 212 ATGATGGCAATTTGTGAAGCCATCTCTTTTCTACTGTCCTTTTCGATGAAGTGACAGAT 153  
QY 6728 AGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCCCTTTGTTGAAAGTCTCAAT 6787  
Db 152 AGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCCCTTTGTTGAAAGTCTCAAT 93  
QY 6788 AGCCCTCTGGTCTCTCGAGACTGTATCTTTGATATCTTTGGAGTAGACGAGAGTGTCTGTG 6847  
Db 92 AGCCCTCTGGTCTCTCGAGACTGTATCTTTGATATCTTTGGAGTAGACGAGAGTGTCTGTG 33  
QY 6848 CTCACCATGTTG 6860  
Db 32 CTCACCATGTTG 20

## RESULT 8

US-09-810-861B-4/c  
; Sequence 4, Application US/09810861B  
; Patent No. 6770799  
; GENERAL INFORMATION:  
; APPLICANT: Mor, Tsafir S.  
; APPLICANT: Soreq, Hermona  
; APPLICANT: Arntzen, Charles J.  
; APPLICANT: Mason, Hugh S.  
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN  
; TITLE OF INVENTION: TRANSGENIC PLANTS  
; FILE REFERENCE: BTI-45  
; CURRENT APPLICATION NUMBER: US/09/810,861B  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/190,440  
; PRIOR FILING DATE: 2000-03-17  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 4  
; LENGTH: 14446  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (11862)..(12157)  
; OTHER INFORMATION: Description of Artificial Sequence: plasmid vector  
; OTHER INFORMATION: pTM036. Identity of sequence residues 11862-12157 unknown.  
US-09-810-861B-4

Query Match 6.1%; Score 461; DB 3; Length 14446;  
Best Local Similarity 86.5%; Pred. No. 5e-80;  
Matches 582; Conservative 0; Mismatches 0; Indels 91; Gaps 3;  
QY 6190 AGATATCATCATCAATCCACTTGTGAAAGACGTGTGGAACGTCCTTTTTCACGA 6249  
Db 14353 AGATATCATCATCAATCCACTTGTGAAAGACGTGTGGAACGTCCTTTTTCACGA 14294  
QY 6250 TGTTCTCTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCA-TCTTGAACG 6308  
Db 14293 TGTTCTCTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCAATCTTGAACG 14234  
QY 6309 ATAGCTTTCTTTATCGCAATGATGGCAATTTGTGAAGCCATCTTCTTTTCTACTGTC 6368  
Db ----- 350



```
Db 14233 ATAGCCTTTCCUATTATCGCAATGATGGCAATTTGTAGAGCCCATCTTCTCTTTCTACTGTC 14174
Qy 6369 CTTTCGATGAAGTGACAGATAGCTCGGCAATCGAATCCGAGGAGTTTCCCGATATTACC 6428
Db 14173 CTTTCGATGAAGTGACAGATAGCTCGGCAATCGAATCCGAGGAGTTTCCCGATATTACC 14114
Qy 6429 CTTTGTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATCTTG 6488
Db 14113 CTTTGTGTGAAAAGT----- 14100
Qy 6489 GAGTAGACGAGAGTGTGCGTCTCCACCATGTTGGGGAATAGATATCAATCAATCCACT 6548
Db 14099 -----TATCAATCAATCCACT 14083
Qy 6549 TGCCTTGAAGACGTGGTGGACGCTTCTTTTCCACGATGTTCTCGTGGGTGGGGT 6608
Db 14082 TGCCTTGAAGACGTGGTGGACGCTTCTTTTCCACGATGTTCTCGTGGGTGGGGT 14023
Qy 6609 CCATCTTTGGGACCACTGTCGGTAGAGCA-TCCTTGAACGATAGCTTTCTTTTATCGCA 6667
Db 14022 CCATCTTTGGGACCACTGTCGGTAGAGCAATCTTGAACGATAGCTTTCTTTATCGCA 13963
Qy 6668 ATGATGGCAATTTGTAGAGCCATCTTCTTTCTACTGTCCTTTCGATGAAGTGACAGAT 6727
Db 13962 ATGATGGCAATTTGTAGAGCCATCTTCTTTCTACTGTCCTTTCGATGAAGTGACAGAT 13903
Qy 6728 AGCTGGGCAATGGAATCCGAGAGGTTCCCGATATTACCTTTGTGAAAAGTCTCAAT 6787
Db 13902 AGCTGGGCAATGGAATCCGAGAGGTTCCCGATATTACCTTTGTGAAAAGTCTCAAT 13843
Qy 6788 AGCCCTCTGGTCTCTGAGACTGTATCTTTGATATCTTGGAGTAGACGAGAGTGTGCGT 6847
Db 13842 AGCCCTCTGGTCTCTGAGACTGTATCTTTGATATCTTGGAGTAGACGAGAGTGTGCGT 13783
Qy 6848 CTCACCAATGTTG 6860
Db 13782 CTCACCAATGTTG 13770

RESULT 9
US-09-737-698B-29/c
; Sequence 29, Application US/09737698B
; Patent No. 6462258
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; TITLE OF INVENTION: No. 6462258el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)C
; CURRENT APPLICATION NUMBER: US/09/737,698B
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/171,173
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 29
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1800)
; OTHER INFORMATION: Y = t/u or c
; OTHER INFORMATION: chimeric promoter fusion CamV and Act8 polynucleotides + Act8 in
; OTHER INFORMATION: tro
US-09-737-698B-29
```

```
Query Match 5.0%; Score 379.8; DB 3; Length 1800;
Best Local Similarity 81.1%; Pred. No. 1.8e-64;
Matches 489; Conservative 0; Mismatches 37; Indels 77; Gaps 1;

Qy 6522 GGGATCTAGATATCAATCAATCCACTTGTGTAAGAGCGTGTGGACGTTCTTTT 6581
Db 535 GGAAGCTTCTATATCAATCAATCCACTTGTGTAAGAGCGTGTGGACGTTCTTTT 476
```

```
Qy 6582 TCCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTTGGGACCACTGTCCGTAGAGCACTCT 6641
Db 475 TCCACGATGTTCTCTGTTGGGTGGGGTCCAATCTTTTGGGACCACTGTCCGTAGAGCACTCT 416
Qy 6642 TGAACGATAGCTTTCTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTTTTCT 6701
Db 415 TCAACGATGGCTTTCTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTTTTCT 356
Qy 6702 ACTGTCTTTTCGATGCAAGTGACAGATAGCTGGGCAATCGAATCCGAGAGGTTTCCCGAT 6761
Db 355 ACTATCTTTCACAATAAAGTGACAGATAGCTGGGCAATCGAATCCGAGAGGTTTCCCGAT 296
Qy 6762 ATTACCTTTCTGTAAGAAAGTCTCAATAGCCCTCTGGTCTTCTTGAGACTGTATCTTTGATA 6821
Db 295 ATTACCTTTCTGTAAGAAAGTCTCA----- 272
Qy 6822 TTTTGGAGTAGACGAGAGTGTGCGTCTCCACCATGTTGGGGAATAGATATCAATCAA 6881
Db 271 -----CATCGGACCAATCAATCAA 253
Qy 6882 TCCACTTCTTTGAAGACGTGGTGGAAAGTCTTCTTTTCCACGATGTTCTCTCGTGGT 6941
Db 252 TCCACTTCTTTGAAGACGTGGTGGAAAGTCTTCTTTTCCACGATGTTCTCTCGTGGT 193
Qy 6942 GGGGTCTCATCTTTTGGGACCACTGTGCGTAGAGGCATCTTGAAACGATAGCCCTTCTTTTA 7001
Db 192 GGGGTCTCATCTTTTGGGACCACTGTGCGTAGAGGCATCTTGAAACGATAGCCCTTCTTTTA 133
Qy 7002 TCGCAATGATGGCAATTTGTAGAGGCATCTTCTTTTCTACTGTCTTTCGATGAAGTGA 7061
Db 132 TCGCAATGATGGCAATTTGTAGAGGCATCTTCTTTTCTACTGTCTTTCGCAATATAAAGTGA 73
Qy 7062 CAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCTTTGTTGAAAAGTC 7121
Db 72 CAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCTTTGTTGAAAAGTC 13
Qy 7122 TCA 7124
Db 12 TCA 10
```

```
RESULT 10
US-09-737-626A-29/c
; Sequence 29, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Plaskinski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6660911el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09/737,626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 29
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1800)
; OTHER INFORMATION: Y = t/u or c
; OTHER INFORMATION: chimeric promoter fusion CamV and Act8 polynucleotides + Act8 in
; OTHER INFORMATION: tro
US-09-737-626A-29
```

```
Query Match 5.0%; Score 379.8; DB 3; Length 1800;
Best Local Similarity 81.1%; Pred. No. 1.8e-64;
Matches 489; Conservative 0; Mismatches 37; Indels 77; Gaps 1;
```

```
QY 6522 GGGATCTAGATATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAAACGTCCTCTTTT 6581
|||
Db 535 GGAAGCTTGATATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAAACGTCCTCTTTT 476
|||
QY 6582 TCCAGATGTTCCCTCGTGGGTGGGGTCCATCTTTTGGGACCACTCTCGGTAGAGGCATCT 6641
|||
Db 475 TCCAGATGCTCCTCGTGGGTGGGGTCCATCTTTTGGGACCACTCTCGGTAGAGGCATCT 416
|||
QY 6642 TGAACGATAGCCCTTTCCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTCT 6701
|||
Db 415 TCAACGATGGCCCTTTCCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTCC 356
|||
QY 6702 ACTGTCCTTTGATGAAGTGAAGATAGTCTGGGCAATGGAATCCGAGAGGTTTCCCGAT 6761
|||
Db 355 ACTATCTTCACAATAAAGTGAAGATAGTCTGGGCAATGGAATCCGAGAGGTTTCCCGAT 296
|||
QY 6762 ATTACCCCTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATA 6821
|||
Db 295 ATTACCCCTTTGTGAAAAGTCTCA----- 272
|||
QY 6822 TTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCAATCAA 6881
|||
Db 271 -----CATCGGACCATCAATCAA 253
|||
QY 6882 TCCACTTGTCTTGAAGACGTGGTTGGAACGTCCTTTTTCACGATGTTCTCTGGGT 6941
|||
Db 252 TCCACTTGTCTTGAAGACGTGGTTGGAACGTCCTTTTTCACGATGTTCTCTGGGT 193
|||
QY 6942 GGGGTCCATCTTTGGGACCACTGTCTGAGGCACTTTGAGGATGCTTTCATGCTCTTTCATGTA 7001
|||
Db 132 TCGCAATGATGTCATTTGTAGAAGCCATCTTCTTCTTCTGCTCTTTCATGTAAGTGA 73
|||
QY 7002 CAGATAGCTGGCAATGGAATCCGAGAGGTTTCCCGATATTTACCCCTTTGTTGAAAAGTC 7121
|||
Db 72 CAGATAGCTGGCAATGGAATCCGAGAGGTTTCCCGATATTTACCCCTTTGTTGAAAAGTC 13
|||
QY 7122 TCA 7124
|||
Db 12 TCA 10

RESULT 11
US-10-427-169-29/c
; Sequence 29, Application US/10427169
; Patent No. 6919495
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flaszinski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6919495el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/10/427,169
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: US/09/737,626A
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1800)
; OTHER INFORMATION: y = t/u or c
; OTHER INFORMATION: chimeric promoter fusion CamV and Act8 polynucleotides + Act8 in
```

```
; OTHER INFORMATION: tiro
US-10-427-169-29
```

```
Query Match 5.0%; Score 379.8; DB 3; Length 1800;
Best Local Similarity 81.1%; Pred. NO. 1.8e-64;
Matches 489; Conservative 0; Mismatches 37; Indels 77; Gaps 1;

QY 6522 GGGATCTAGATATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAAACGTCCTCTTTT 6581
|||
Db 535 GGAAGCTTGATATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAAACGTCCTCTTTT 476
|||
QY 6582 TCCAGATGTTCCCTCGTGGGTGGGGTCCATCTTTTGGGACCACTCTCGGTAGAGGCATCT 6641
|||
Db 475 TCCAGATGCTCCTCGTGGGTGGGGTCCATCTTTTGGGACCACTCTCGGTAGAGGCATCT 416
|||
QY 6642 TGAACGATAGCCCTTTCCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTCT 6701
|||
Db 415 TCAACGATGGCCCTTTCCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTCC 356
|||
QY 6702 ACTGTCCTTTGATGAAGTGAAGATAGTCTGGGCAATGGAATCCGAGAGGTTTCCCGAT 6761
|||
Db 355 ACTATCTTCACAATAAAGTGAAGATAGTCTGGGCAATGGAATCCGAGAGGTTTCCCGAT 296
|||
QY 6762 ATTACCCCTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATA 6821
|||
Db 295 ATTACCCCTTTGTGAAAAGTCTCA----- 272
|||
QY 6822 TTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCAATCAA 6881
|||
Db 271 -----CATCGGACCATCAATCAA 253
|||
QY 6882 TCCACTTGTCTTGAAGACGTGGTTGGAACGTCCTTTTTCACGATGTTCTCTGGGT 6941
|||
Db 252 TCCACTTGTCTTGAAGACGTGGTTGGAACGTCCTTTTTCACGATGTTCTCTGGGT 193
|||
QY 6942 GGGGTCCATCTTTGGGACCACTGTCTGAGGCACTTTGAGGATGCTTTCATGCTCTTTCATGTA 7001
|||
Db 132 TCGCAATGATGTCATTTGTAGAAGCCATCTTCTTCTTCTGCTCTTTCATGTAAGTGA 73
|||
QY 7002 CAGATAGCTGGCAATGGAATCCGAGAGGTTTCCCGATATTTACCCCTTTGTTGAAAAGTC 7121
|||
Db 72 CAGATAGCTGGCAATGGAATCCGAGAGGTTTCCCGATATTTACCCCTTTGTTGAAAAGTC 13
|||
QY 7122 TCA 7124
|||
Db 12 TCA 10
```

```
RESULT 12
US-10-427-180-29/c
; Sequence 29, Application US/10427180
; Patent No. 6949696
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen
; APPLICANT: Flaszinski, Stanislaw
; APPLICANT: Wilkinson, Jack
; TITLE OF INVENTION: No. 6949696el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/10/427,180
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: US/09/737,626A
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 1800
; TYPE: DNA
```

```
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1800)
; OTHER INFORMATION: y = t/u or c
; OTHER INFORMATION: chimeric promoter fusion CamV and Act8 polynucleotides + Act8 in
; OTHER INFORMATION: tro
US-10-427-180-29

Query Match      5.0%; Score 379.8; DB 3; Length 1800;
Best Local Similarity 81.1%; Pred. No. 1.8e-64;
Matches 489; Conservative 0; Mismatches 37; Indels 77; Gaps 1;

Qy 6522 GGGATCTAGATATACATCAATCCACTTGCCTTTGAAGACGCTGTTGGAACGCTCTCTTTT 6581
Db 535 GGAAGCTTGATATCAATCAATCCACTTGCCTTTGAAGACGCTGTTGGAACGCTCTCTTTT 476

Qy 6582 TCCACGATGTTCTCTCGTGGGTGGGGTCCATCTTTTGGGACCACTGTCCGTAGAGGCATCT 6641
Db 475 TCCACGATGCTCCTCGTGGGTGGGGTCCATCTTTTGGGACCACTGTCCGTAGAGGCATCT 416

Qy 6642 TGAACGATAGCCCTTTCTTTATCGCAATGATGGCAATTTGTAGAACCACTCTTCTTTTCT 6701
Db 415 TCAACGATGGCCCTTTCTTTATCGCAATGATGGCAATTTGTAGAACCACTCTTCTTTTCT 356

Qy 6702 ACTGTCCTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGAT 6761
Db 355 ACTATCTTTCACAAATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGAT 296

Qy 6762 ATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTGATA 6821
Db 295 ATTACCCCTTTGTTGAAAAGTCTCA----- 272

Qy 6822 TTCTTGGAGTAGACGAGAGTGTCGTCTCCACCATGTTGGGATCTAGATATCACATCAA 6881
Db 271 -----CATCGGCATCACATCAA 253

Qy 6882 TCCACTTGTGTAAGAGCGTGGTTGGAACGCTCTCTTTTCCACGATGTTCTCTGTGGT 6941
Db 252 TCCACTTGTGTAAGAGCGTGGTTGGAACGCTCTCTTTTCCACGATGTTCTCTGTGGT 193

Qy 6942 GGGGGTCCATCTTGGGACCACTGTCGGTAGAGGCATCTTGAAACGATAGCCCTTCTCTTA 7001
Db 192 GGGGGTCCATCTTGGGACCACTGTCGGTAGAGGCATCTTCAACGATGGCTTCTCTTA 133

Qy 7002 TCGCAATGATGGCAATTTGTAGAACCATCTTCTTTTCTACTGTCTTTCGATGAAGTGA 7061
Db 132 TCGCAATGATGGCAATTTGTAGAACCATCTTCTTTTCTACTGTCTTTCGATGAAGTGA 73

Qy 7062 CAGATAGCTGGGCAATGGNAATCCGAGAGGTTTCCGATATTTACCTTTGTTGAAAAGTC 7121
Db 72 CAGATAGCTGGGCAATGGNAATCCGAGAGGTTTCCGATATTTACCTTTGTTGAAAAGTC 13

Qy 7122 TCA 7124
Db 12 TCA 10
```

```
RESULT 13
US-09-737-698B-30/c
; Sequence 30, Application US/09737698B
; Patent No. 6462258
; GENERAL INFORMATION:
; APPLICANT: Wilkison, Jack
; APPLICANT: Finchner, Karen
; TITLE OF INVENTION: No. 6462258el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)C
; CURRENT APPLICATION NUMBER: US/09/737,698B
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/171,173
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 30
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1742)
; OTHER INFORMATION: chimeric promoter fusion CamV and Act2 polynucleotides + Act2 in
; OTHER INFORMATION: tro
US-09-737-698B-30

Query Match      5.0%; Score 379.2; DB 3; Length 1742;
Best Local Similarity 81.0%; Pred. No. 2.4e-64;
Matches 489; Conservative 0; Mismatches 38; Indels 77; Gaps 1;

Qy 6182 TCGGATCTAGATATACATCAATCCACTTGCCTTTGAAGACGCTGTTGGAACGCTCTCTTTT 6241
Db 535 TTGAAGCTTGATATACATCAATCCACTTGCCTTTGAAGACGCTGTTGGAACGCTCTCTTTT 476

Qy 6242 TTCCACGATGTTCTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCCGTAGAGGCATC 6301
Db 475 TTCCACGATGCTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCCGTAGAGGCATC 416

Qy 6302 TTGAACGATAGCCCTTTCTTTATCGCAATGATGGCAATTTGTAGAACCACTTCTCTTTTCT 6361
Db 415 TTCAACGATGGCCCTTTCTTTATCGCAATGATGGCAATTTGTAGAGGCACCTTCTCTTTTCT 356

Qy 6362 TACTGTCTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 6421
Db 355 CACTATCTTCAACAATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 296

Qy 6422 TATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTGAT 6481
Db 295 TATTACCCCTTTGTTGAAAAGTCTCA----- 271

Qy 6482 ATTCTTGGAGTAGACGAGAGTGTCGTCTCCACCATGTTGGGATCTAGATATCACATCA 6541
Db 270 -----CATCGGCATCACATCA 253

Qy 6542 ATCCACTTGTGTAAGACGCTGGTTGGAACGCTCTCTTTTCCACGATGTTCTCTGTGG 6601
Db 252 ATCCACTTGTGTAAGACGCTGGTTGGAACGCTCTCTTTTCCACGATGTTCTCTGTGG 193

Qy 6602 TGGGGTCCATCTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCCTTCTCTTTT 6661
Db 192 TGGGGTCCATCTTGGGACCACTGTCGGTAGAGGCATCTTCAACGATAGCCCTTCTCTTTT 133

Qy 6662 ATCGCAATGATGGCAATTTGTAGAACCATCTTCTTTTCTACTGTCTTTCGATGAAGTG 6721
Db 132 ATCGCAATGATGGCAATTTGTAGAACCATCTTCTTTTCTACTGTCTTTCGATGAAGTG 73

Qy 6722 ACAGATAGCTGGGCAATGGNAATCCGAGAGGTTTCCGATATTTACCTTTGTTGAAAAGT 6781
Db 72 ACAGATAGCTGGGCAATGGNAATCCGAGAGGTTTCCGATATTTACCTTTGTTGAAAAGT 13

Qy 6782 CTC 6785
Db 12 CTC 9
```

```
RESULT 14
US-09-737-626A-30/c
; Sequence 30, Application US/09737626A
; Patent No. 6660911
; GENERAL INFORMATION:
; APPLICANT: Finchner, Karen
; APPLICANT: Wilkison, Jack
; APPLICANT: Stanislav
; TITLE OF INVENTION: No. 6660911el Plant Expression Constructs
; FILE REFERENCE: 38-21(51499)B
; CURRENT APPLICATION NUMBER: US/09/737,626A
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/737,626
```

; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 30  
; LENGTH: 1742  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (1)..(1742)  
; OTHER INFORMATION: chimeric promoter fusion CamV and Act2 in  
; OTHER INFORMATION: tro  
US-09-737-626A-30

Query Match 5.0%; Score 379.2; DB 3; Length 1742;  
Best Local Similarity 81.0%; Pred. No. 2.4e-64;  
Matches 489; Conservative 0; Mismatches 38; Indels 77; Gaps 1;  
QY 6182 TCGGATCTAGATATCACATCAATCCACTTTCCTTTGAGACGTTGGAACTCTCTTT 6241  
Db |||||  
QY 535 TTGAAGCTTGTATATCACATCAATCCACTTTCCTTTGAGACGTTGGAACTCTCTTT 476  
Db |||||  
QY 6242 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATC 6301  
Db |||||  
QY 475 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATC 416  
Db |||||  
QY 6302 TTGAACGATGAGCTTTCTTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 6361  
Db |||||  
QY 415 TTCAACGATGGCTTTCTTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 356  
Db |||||  
QY 6362 TACTGTCTTTGATGAAGTGCAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 6421  
Db |||||  
QY 355 CACTATCTTCAATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 296  
Db |||||  
QY 6422 TATTACCTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTCTGAGACTGTATCTTTGAT 6481  
Db |||||  
QY 295 TATTACCTTTGTGAAAAGTCTCA----- 271  
Db |||||  
QY 6482 ATTTCTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCATCA 6541  
Db |||||  
QY 270 -----CATCGGACCATCATCA 253  
Db |||||  
QY 6542 ATCCACTTGTGTAAGACGTTGGTAACGTTCTTTTCCAGATGTTCTCGTGG 6601  
Db |||||  
QY 252 ATCCACTTGTGTAAGACGTTGGTAACGTTCTTTTCCAGATGTTCTCGTGG 193  
Db |||||  
QY 6602 TGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTCAACGATGTTCTCTTT 6661  
Db |||||  
QY 192 TGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTCAACGATGTTCTCTTT 133  
Db |||||  
QY 6662 ATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCTTTTCGATGAAGTG 6721  
Db |||||  
QY 132 ATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCTTTTCGATGAAGTG 73  
Db |||||  
QY 6722 ACAGATAGCTGGCAATGGAATCCGAGAGGTTTCCGATATATACCTTTGTGAAAAGT 6781  
Db |||||  
QY 72 ACAGATAGCTGGCAATGGAATCCGAGAGGTTTCCGATATATACCTTTGTGAAAAGT 13  
Db |||||  
QY 6782 CTCA 6785  
Db |||||  
QY 12 CTCA 9

## RESULT 15

US-10-427-169-30/C  
; Sequence 30, Application US/10427169  
; Patent No. 6919495  
; GENERAL INFORMATION:  
; APPLICANT: Fincher, Karen  
; APPLICANT: Flaszinski, Stanislaw  
; APPLICANT: Wilkinson, Jack  
; TITLE OF INVENTION: No. 6919495el Plant Expression Constructs  
; FILE REFERENCE: 38-21(51499)B

; CURRENT APPLICATION NUMBER: US/10/427,169  
; CURRENT FILING DATE: 2003-05-01  
; PRIOR APPLICATION NUMBER: US/09/737,626A  
; PRIOR FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 09/737,626  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 30  
; LENGTH: 1742  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (1)..(1742)  
; OTHER INFORMATION: chimeric promoter fusion CamV and Act2 in  
; OTHER INFORMATION: tro  
US-10-427-169-30

Query Match 5.0%; Score 379.2; DB 3; Length 1742;  
Best Local Similarity 81.0%; Pred. No. 2.4e-64;  
Matches 489; Conservative 0; Mismatches 38; Indels 77; Gaps 1;  
QY 6182 TCGGATCTAGATATCACATCAATCCACTTTCCTTTGAGACGTTGGAACTCTCTTT 6241  
Db |||||  
QY 535 TTGAAGCTTGTATATCACATCAATCCACTTTCCTTTGAGACGTTGGAACTCTCTTT 476  
Db |||||  
QY 6242 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATC 6301  
Db |||||  
QY 475 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATC 416  
Db |||||  
QY 6302 TTGAACGATGAGCTTTCTTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 6361  
Db |||||  
QY 415 TTCAACGATGGCTTTCTTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 356  
Db |||||  
QY 6362 TACTGTCTTTGATGAAGTGCAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 6421  
Db |||||  
QY 355 CACTATCTTCAATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 296  
Db |||||  
QY 6422 TATTACCTTTGTGAAAAGTCTCAATAGCCCTCTGGTCTCTGAGACTGTATCTTTGAT 6481  
Db |||||  
QY 295 TATTACCTTTGTGAAAAGTCTCA----- 271  
Db |||||  
QY 6482 ATTTCTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCATCA 6541  
Db |||||  
QY 270 -----CATCGGACCATCATCA 253  
Db |||||  
QY 6542 ATCCACTTGTGTAAGACGTTGGTAACGTTCTTTTCCAGATGTTCTCGTGG 6601  
Db |||||  
QY 252 ATCCACTTGTGTAAGACGTTGGTAACGTTCTTTTCCAGATGTTCTCGTGG 193  
Db |||||  
QY 6602 TGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTCAACGATGTTCTCTTT 6661  
Db |||||  
QY 192 TGGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTCAACGATGTTCTCTTT 133  
Db |||||  
QY 6662 ATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCTTTTCGATGAAGTG 6721  
Db |||||  
QY 132 ATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCTTTTCGATGAAGTG 73  
Db |||||  
QY 6722 ACAGATAGCTGGCAATGGAATCCGAGAGGTTTCCGATATATACCTTTGTGAAAAGT 6781  
Db |||||  
QY 72 ACAGATAGCTGGCAATGGAATCCGAGAGGTTTCCGATATATACCTTTGTGAAAAGT 13  
Db |||||  
QY 6782 CTCA 6785  
Db |||||  
QY 12 CTCA 9

Search completed: December 31, 2005, 05:24:05  
Job time : 851 secs

**BEST AVAILABLE COPY**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2005, 15:23:45 ; Search time 17827 Seconds  
(without alignments)  
19893.761 Million cell updates/sec

Title: US-10-650-249-1  
Perfect score: 7580  
Sequence: 1 agctctataattcaagaga.....gccgcaccgcgggtggaagct 7580

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gss1:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1256.4	16.6	1466	4	CNS0A60U
2	1118	14.7	1371	4	US-10-650-249-1
3	964.8	12.7	1320	4	CNS0A50A
4	720.4	9.5	736	10	CNS0A50A
5	711	9.4	720	10	CNS0A50A
6	688.4	9.1	972	8	DR750042
7	672.2	8.9	868	8	DR749973
8	671.4	8.9	866	8	DR750041
9	570.2	7.5	733	8	CX308688
10	528	7.0	528	10	CNS00PNT
11	477.4	6.3	597	10	CG729372
12	461.2	6.1	486	9	B26529
13	442	5.8	452	9	AQ011657
14	431.4	5.7	433	9	B97814
15	424.8	5.6	468	10	CL517384
16	412.2	5.4	690	7	CF854584
17	381.2	5.0	655	7	CK437920
18	381.2	5.0	698	7	CK438348
19	381.2	5.0	698	7	CK438483
20	381.2	5.0	702	7	CO472768
21	381.2	5.0	713	7	CK437984
22	381.2	5.0	736	7	CK438485

C 23	381.2	5.0	741	7	CK438362
C 24	381.2	5.0	755	7	CK438332
C 25	381.2	5.0	803	7	CO472761
C 26	371	4.9	446	1	AA585745
C 27	350.2	4.6	492	9	AQ361936
C 28	336.2	4.4	839	9	AQ362115
C 29	324.6	4.3	441	7	CK437874
C 30	319.2	4.2	869	10	CW799882
C 31	315	4.2	610	7	CO481168
C 32	310	4.1	654	10	CG206846
C 33	305.4	4.0	507	8	CV961115
C 34	305.4	4.0	557	8	CV965337
C 35	305.4	4.0	609	8	CV965230
C 36	305.4	4.0	692	8	CV964956
C 37	304.6	4.0	490	9	BZ586494
C 38	304.6	4.0	534	9	BZ586497
C 39	304.6	4.0	570	9	BZ586953
C 40	304.6	4.0	741	10	CZ908574
C 41	304.4	4.0	643	10	CG805032
C 42	303.8	4.0	431	9	BH753412
C 43	303.6	4.0	385	9	BH212594
C 44	303.6	4.0	435	9	BH171104
C 45	303.6	4.0	438	9	BH619277

## ALIGNMENTS

RESULT 1  
LOCUS CNS0A60U 1466 bp mRNA linear HTC 06-FEB-2004  
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUFB52B04 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).  
ACCESSION BX822604  
VERSION BX822604.1  
KEYWORDS HTC; GSUFB52B04  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 1466)  
AUTHORS Castell, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1466)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT The sequences are based on single pass reads.  
Life technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
URV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full-length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES  
source  
1..1466  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"

		/db_xref="taxon:3702"			
		/clone="GSLTFB522B04"			
		/tissue_type="Flowers and buds"			
		/ecotype="Col-0"			
		/plasmid="pCMVSPORT_6"			
		1. .1466			
		/gene="At3g55370"			
ORIGIN					
Query Match 16.6%; Score 1256.4; DB 4; Length 1466;					
Best Local Similarity 99.1%; Pred. No. 3.6e-213;					
Matches 1274; Conservative 0; Mismatches 11; Indels 1; Gaps 1;					
QY	4105	GCAGCAAGGGAACCAACATCAGCTAGAAATGTGTCAAACTGACCAAGAACCCCTAATAATTA	4164		
DB	173	GCAGCAAGGGAACCAACATCAGCTAGAAATGTGTTCACACTTACCAAGAACCCCTAATAATTA	232		
QY	4165	CTTACGGCAGCTCTCATCAACCGACTTCTCAGGTTGCAGGTTTCGAGTCAAGCTAGAGT	4224		
DB	233	CTTACGGCAGCTCTCATCAACCGACTTCTCAGGTTGCAGGTTTCGAGTCAAGCTAGAGT	292		
QY	4225	GAATTCATGTTGGAACTGCTCGGATCGCAAAAGTCCCATGCTGAAGCAGCTCTAAA	4284		
DB	293	CAATTCATGTTGGAACTGCTCGGATCGCAAAATCCCATTTGCCCTGAAGCAGCTCTAAA	352		
QY	4285	TTGCCCTAGATGTGACTCAACCAATCTAAAGTTCTGTACTTCAATAACTATAGCCTTAC	4344		
DB	353	TTGCCCTAGATGTGACTCAACCAATCTAAAGTTCTGTACTTCAATAACTATAGCCTTAC	412		
QY	4345	TCAACTCGCCATTTCTGCAAAACATGTCGTCGCTATTGGACACGCTGGCGGTTCTTTGAG	4404		
DB	413	TCAACTCGCCATTTCTGCAAAACATGTCGTCGCTATTGGACACGCTGGCGGTTCTTTGAG	472		
QY	4405	GAATGTTCTGTGTGAGGAGGCTTTAGGAGGAACAAGAGACCAATCCAGATCGAAATC	4464		
DB	473	GAATGTTCTGTGTGAGGAGGCTTTAGGAGGAACAAGATAGCAAAATCCAGATCGAAATC	532		
QY	4465	TACGCTCGTGCTCTCGACTGATAATACTACTAGTACTTCACTACTTCTTCGCCCAAG	4524		
DB	533	TACGCTCGTGCTCTCGACTGATAATACTACTAGTACTTCACTACTTCTTCGCCCAAG	592		
QY	4525	TTACTCAAACCCCTAGCAAGTTTCTATAGCTACGGTCAAAATCCCGGAGTTTAATTCCAACTT	4584		
DB	593	TTACTCAAACCCCTAGCAAGTTTCTATAGCTACGGTCAAAATCCCGGAGTTTAATTCCAACTT	652		
QY	4585	GCCCATCTTGCTCTCTCCAAAGCCTTGGAGATTACAAATTCAGCAACAACCTGGATTAGA	4644		
DB	653	GCCCATCTTGCTCTCTCCAAAGCCTTGGAGATTACAAATTCAGCAACAACCTGGATTAGA	712		
QY	4645	TTTTTGTGGAACTCAAAATAGCAACATGATAGTGTGATGATTTCTAGTGGTGGGATCTT	4704		
DB	713	TTTTTGTGGAACTCAAAATAGCAACATGATAGTGTGATGATTTCTAGTGGTGGGATCTT	772		
QY	4705	GGATGATGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTTGATCAACAC	4764		
DB	773	GGATGATGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTTGATCAACAC	832		
QY	4765	TACCGAATTTGGTGCATCTTCAAACCGGTTATATCATTTACTAGAGGTAAGGAGGTGT	4824		
DB	833	TACCGAATTTGGTGCATCTTCAAACCGGTTATATCATTTACTAGAGGTAAGGAGGTGT	892		
QY	4825	TAATCAAGGTGATTTCTCAACAGAGTAGTGATTTATTCCAATCAGCTTAATGTTTAAGCC	4884		
DB	893	TAATCAAGGTGATTTCTCAACAGAGTAGTGATTTATTCCAATCAGCTTAATGTTTAAGCC	952		
QY	4885	CTTGATGGATTTTCTTTCAGGCGGGGTTAGCCGCCACGCAACAAGAAATGTGAAGCGGGA	4944		
DB	953	CTTGATGGATTTTCTTTCAGGCGGGGTTAGCCGCCACGCAACAAGAAATGTGAAGCGGGA	1012		
QY	4945	AGAGAATGATCAGGATCGGGGTAGGGATCGGGATGGAGTGAATACTTATCAAGAAACTT	5004		
DB	1013	AGAGAATGATCAGGATCGGGGTAGGGATCGGGATGGAGTGAATACTTATCAAGAAACTT	1072		
QY	5005	TTTGGGTAATATCAACATAAACTCAGGCAGGAACGAGGAATACATCATCGGGAGGTAA	5064		
DB	1073	TTTGGGTAATATCAACATAAACTCAGGCAGGAACGAGGAATACATCATCGGGAGGTAA	1132		
QY	5065	CAGTTCTTGGACCGGTTTCACTCTCAACAACTCAACAGGCCATCTCTCATTTCTAAGTACT	5124		
DB	1133	CAGTTCTTGGACC-GTTTCACCTCCAACTCAACAGGCCATCTCTCATTTCTAAGTACT	1191		
QY	5125	CAGCACTAGCTATTCTTGATGATTTCTTTTGTGGTGGGGTGTACATTTGGTGTCTTGTCTAT	5184		
DB	1192	CAGCACTAGCTATTCTTGATGATTTCTTTTGTGGTGGGGTGTACATTTGGTGTCTTGTCTAT	1251		
QY	5185	CGGAGTTATTCTCTGAGGAAGATCAAAACCATCAGCTATATCCAAAGGCTAAATTTGAGGC	5244		
DB	1252	CGGAGTTATTCTCTGAGGAAGATCAAAACCATCAGCTATATCCAAAGGCTAAATTTGAGGC	1311		
QY	5245	TCAAAGGAAAGGTATGGTTATAAAACATCTTTTGTGATCTTTTAAAGATCTTCAAAGTG	5304		
DB	1312	TCAAAGGAAAGGTATGGTTATAAAACATCTTTTGTGATCTTTTAAAGATCTTCAAAGTG	1371		
QY	5305	TGAGTATGTTTATTGGTTGGCTTCTGGTGATATTATTTTATTAGAAATTTGGTCTTAT	5364		
DB	1372	TGAGTATGTTTATTGGTTGGCTTCTGGTGATATTATTTTATTAGAAATTTGGTCTTAT	1431		
QY	5365	ATATTGGCTATATATAGAGGTGTGGG	5390		
DB	1432	ATATTGGCTATATATAGAGGTGTGGG	1457		
RESULT 2					
CNS0A5T8					
LOCUS					
DEFINITION					
Arabidopsis thaliana Full-length cDNA Complete sequence from clone					
GSLTFB532B03 of Flowers and buds of strain col-0 of Arabidopsis					
thaliana (thale cress).					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
JOURNAL					
COMMENT					
The sequences are based on single pass reads.					
Life Technologies (a division of Invitrogen)					
full-length libraries construction : Temple G.					
Genoscope members carried out sequencing and annotation : Castelli					
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,					
Schachter V., Weissenbach J., Salanoubat M.					
URV INRA : Clepet C., Caboche M.					
Annotation is based on the June 2003 version of the Arabidopsis					
genome released by MIPS (Munich Information center for Protein					
Sequences). 5 prime and 3 prime are assembled with Phrap.					
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full					
length					
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.					
Location/Qualifiers					
1. .1371					
/organism="Arabidopsis thaliana"					
FEATURES					
source					



```
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTF532B03"
/tissue_type="Flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..1371
/gene="At3g55370"

ORIGIN
Query Match          14.7%; Score 1118; DB 4; Length 1371;
Best Local Similarity 92.9%; Pred. No. 1.5e-188;
Matches 1221; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

QY 4105 GCAGCAAGGAAACCAACATCAGCTAGAATGTCTCAAACTGACACGAGAACCCCTAATAATTA 4164
DB 151 GCAGCAAGGAAACCAACATCAGCTAGAATGTCTCAAACTGACACGAGAACCCCTAATAATTA 210
QY 4165 CTTACGGCAGCTCTCATCACCAACGACTTCTCAGGTTCAGGTTGCAGGTTTCGAGTCAAGCTAGAGT 4224
DB 211 CTTACGGCAGCTCTCATCACCAACGACTTCTCAGGTTCAGGTTTCGAGTCAAGCTAGAGT 270
QY 4225 GAATTCATAGTGGGAAGCTGCTCGATCGCAAAAGTCCCATTCGCTGAAGCAGCTCTAAA 4284
DB 271 GAATTCATAGTGGGAAGCTGCTCGATCGCAAAAGTCCCATTCGCTGAAGCAGCTCTAAA 330
QY 4285 TTGCCCTAGATGTGACTCAACCAATACCTAAGTTCTGTCTTCACTCAATCACTATAGCCTTAC 4344
DB 331 TTGCCCTAGATGTGACTCAACCAATACCTAAGTTCTGTCTTCACTCAATCACTATAGCCTTAC 390
QY 4345 TCAACTCGCCATTTCTGCAAAACATGCTCGCTATTGGACACGTCGGGGTTCCTTGAG 4404
DB 391 TCAACTCGCCATTTCTGCAAAACATGCTCGCTATTGGACACGTCGGGGTTCCTTGAG 450
QY 4405 GAATGTTCTGTGTGGAGGAGCTTTAGGAGGAACAGAGCAATCCAGATCGAATC 4464
DB 451 GAATGTTCTGTGTGGAGGAGCTTTAGGAGGAACAGAGCAATCCAGATCGAATC 510
QY 4465 TACGGTCTGGTCTCGACTGATAATCTACTAGTACTTTCATCATTACTTCTCGGCCAAG 4524
DB 511 TACGGTCTGGTCTCGACTGATAATCTACTAGTACTTTCATCATTACTTCTCGGCCAAG 570
QY 4525 TTACTCAAAACCTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATTCACACTT 4584
DB 571 TTACTCAAAACCTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATTCACACTT 630
QY 4585 GCCCATCTGCTCTCTCCAAAGCTTGGAGATTACAAATTCAGCAACACCTGGATTAGA 4644
DB 631 GCCCATCTGCTCTCTCCAAAGCTTGGAGATTACAAATTCAGCAACACCTGGATTAGA 690
QY 4645 TTTTGGTGGAACTCAAAATAGCAACATGATAGTGTATGAGTTCTAGTGGTGGGATCTT 4704
DB 691 TTTTGGTGGAACTCAAAATAGCAACATGATAGTGTATGAGTTCTAGTGGTGGGATCTT 750
QY 4705 GGATGCATGGAGAATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTGATCAACAC 4764
DB 751 GGATGCATGGAGAATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTGATCAACAC 810
QY 4765 TACCGGATTGGTGCAATCTTCAACGCGTTATATCCATTACTAGAGGTAAGGAGGTGT 4824
DB 811 TACCGGATTGGTGCAATCTTCAACGCGTTATATCCATTACTAGAGGTAAGGAGGTGT 855
QY 4825 TAATCAAGTGTATCTCAACAGAGTAGTATTATTCATCAATCAGCTAATGTTTAAAGCC 4884
DB 856 ----- 855
QY 4885 CTTGTGATGATTTTCTTCAGGGGGGTTAGCCGACGCAACAAAGAAATGTGAAGCGGA 4944
DB 856 -----AGGGGGGTTAGCCGACGCAACAAAGAAATGTGAAGCGGA 897
QY 4945 AGAAGATGATCAGGATCGGGGTAGGGATGGGATGGGATGGGATGGAATACTTATCAAGAACTT 5004
DB 898 AGAAGATGATCAGGATCGGGGTAGGGATGGGATGGGATGGGATGGAATACTTATCAAGAACTT 957
```

```
QY 5005 TTTGGGTAATATCAACATAAACTCAGGCAGGAAACGAGGAATACATCATCGGGAGGTAA 5064
DB 958 TTTGGGTAATATCAACATAAACTCAGGCAGGAAACGAGGAATACATCATCGGGAGGTAA 1017
QY 5065 CAGTTCCTTGGACCGGTTTCACTCCAACTCAACAGGCCATCTCTCATTTAAGTACT 5124
DB 1018 CAGTTCCTTGGACCGGTTTCACTCCAACTCAACAGGCCATCTCTCATTTAAGTACT 1077
QY 5125 CAGCACTAGCTATTCTTCGATGATCTTTTGGTGGTGGGGTGTACATTTGGTCTGTGAT 5184
DB 1078 CAGCACTAGCTATTCTTCGATGATCTTTTGGTGGTGGGGTGTACATTTGGTCTGTGAT 1137
QY 5185 GCGAGTTATTCTGAGGAAGATCAAACTCAGCTATATCAAAAGGCTAAATTTTGGAGC 5244
DB 1138 GCGAGTTATTCTGAGGAAGATCAAACTCAGCTATATCAAAAGGCTAAATTTTGGAGC 1197
QY 5245 TCAAAGGAAAGTATGTTTATAAACTATCTTTTGTGATCTTTTAAAGATCTTCAAAGTG 5304
DB 1198 TCAAAGGAAAGTATGTTTATAAACTATCTTTTGTGATCTTTTAAAGATCTTCAAAGTG 1257
QY 5305 TGAATATGTTTATTGGTGGCTTCTGGTGAATTTATGTTTATTAGAAATTTGGTCTTAT 5364
DB 1258 TGAATATGTTTATTGGTGGCTTCTGGTGAATTTATGTTTATTAGAAATTTGGTCTTAT 1317
QY 5365 ATATTGGCTATATAGAGGTGCTGGTGAATTTATGATTAATTTCAAGAGTTGATGT 5418
DB 1318 ATATTGGCTATATAGAGGTGCTGGTGAATTTATGATTAATTTCAAGAGTTGATGT 1371

RESULT 3
CNS0A504
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTF532B02 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
BX824803.1 GI:42464248
VERSION
HTC; GSLT cDNA.
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 1320)
Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 1320)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1..1320
FEATURES
source
```

gene		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
		/clone="GSLTPGH73ZF02"	
		/tissue_type="Hormone Treated Callus"	
		/ecotype="Col-0"	
		/plasmid="pCMVSPORT_6"	
1..1320			
/gene="At3g55370"			
ORIGIN			
Query Match		12.7%; Score 964.8; DB 4; Length 1320;	
Best Local Similarity		85.9%; Pred. No. 2.9e-161;	
Matches 1134; Conservative		0; Mismatches 92; Indels 94; Gaps 2;	
QY	4105	GCAGCAAGGGAACCAACATCAGCTAGATGTGTCAAACTGACGACGAGACCCCTAATAATTA	4164
Db	95	GCAGCGAGGGAACGAACATGAGCGAGGGTGTGTGAGAACTGAGCAGGACCCCTAATGATTA	154
QY	4165	CTTACGGCAGCTCTCATCACCACGACTTCTCAGGTTTCAGGTTTCGAGTCAAGCTAGAGT	4224
Db	155	CTTACGGCAGCGCTGATCACCACGACTTCTCAGGTTTCAGGTTTCGAGTTCGAGTGAAGCGAGAGT	214
QY	4225	GAATTCATGTTGGTGAACGTGCTCGATCGCAAAAGTCCCATTTGCCCTGAAGCAGCTCTAAA	4284
Db	215	GAGTTTCGATGGTGAACGTGCTCGATCGCGAGAGTTCATTTGCTGAAGCAGCGCGGAG	274
QY	4285	TTGCCCTAGATGTGACTCAACCAATACTAGTTCTGTTACTTTCGAATACTATAGCCCTTAC	4344
Db	275	TTGCCCTAGATGTGACTGAACCAATACTAGGAGTTCTGTTACTCAATGACGATAGCCCTTAC	334
QY	4345	TCAACTCGCCATTTCTGCAAAACATGCTGCTGCTATTGGACACGTCGGCGGTTCTCTTGAG	4404
Db	335	TCAACTCGCCATTTCTGCGAAGATGTCGTCGCTATTGGACACGTCGGCGGTTCTCTTGAG	394
QY	4405	GAATGTTCTGTTGGAGAGGCTTTAGGAGGAACAAGAGAAACAATCCAGATCGAAATC	4464
Db	395	GAATGTTCTGTTGGAGAGGCTTTAGGAGGAACAAGATAAGCGAATCGAGATCGAGATC	454
QY	4465	TACGGTCGTGGTCTCGACTGATATACTACTAGTACTTTCATCATCTTCTTCGCCCAAG	4524
Db	455	TACGGTCGTGGTCTCGACTGAGAGAGACTACGAGTACTTTCATCTACTTTTTCGCCCAAG	514
QY	4525	TTACTCAAACTTAGCAAGTTTCATAGTACGGTCAAAATCCCGGAGTTTAAATCCCAACTT	4584
Db	515	TGACTCAAACTTAGCGAGTGTCTATAGGACGGTCAAAATCCCGGAGTCCCAATTCGACTT	574
QY	4585	GCCCATCTTGCTCTCTCTCAAAGCTTTGGAGATTACAATTCGAAGCAACACTGGATTAGA	4644
Db	575	GCCCATCTTGCTCTCTCTCAAAGCTTTGGAGATTACAATTCGAAGCAACACTGGATTAGA	634
QY	4645	TTTTGGTGGAACTCAAAATGAACCAACATGATAGTGTATGAGTTCTAGTGGTGGGATCTT	4704
Db	635	TTTTGGTGGAACTAACAATACGACATGATAAGTGGTATGAGTCGAGTCGAGTGGGATCTC	694
QY	4705	GGATGCAATGGAGAACTCTCCATCACAACAGCTCAGCAATTCCTCTTTTGATCAACAC	4764
Db	695	GGATGCAATGGAGAACTCTCTGATGCAACAAGCTCAGCAATTCCTCTCTTGATCAACAC	754
QY	4765	TACCGGATTGGTGCAATCTTCAAAACGGGTATATCCATTACTAGAAGGTAAAGGAGGTGT	4824
Db	755	TACCGCATTTGGTGCAATCTTCAAAACGGGTATATCCATTACTCGA-----	799
QY	4825	TAATCAAGGTGATTCTCAACAGAGAGTAGTGTATTTCCAATCAGCTAAATGTTTAAGCC	4884
Db	800	-----	799
QY	4885	CTTGATGGATTTTCTTCAGCGCGGGTTAGCGCCAGCAACAAAGAAATGTGAAGCGGGA	4944
Db	800	-----AGCGGGGTTAGCGCCAGCAACAAAGAAATGTGAAGCGGGA	841
QY	4945	AGAGAAATGATCAGGATCGGGTATGGGATGGGGATGGAGTGAATACTTATCAAGAAACTT	5004

Db	842	AGAGATGATCAGGATCGGGTAGGGATGGGATGGAGTGAATTAACCTTATCAAGAAACTC	901
Qy	5005	TTTGGGTAATATCAACATAAACTCAGGCAGGAAACGAGGAATACACATCATCGGGAGGTAA	5064
Db	902	CTTGGGTAATATCAACATAAACTCAGGCAGGAAACGAGGAATACACATCATCGGGAGGTAA	961
Qy	5065	CAGTCTTTGGACCGGTTTACCTCCAACTCAACAGGCCATCTCTCATCTTAAGTACT	5124
Db	962	CAGTCTTTGGACC-GTTTACCTCCAACTCAACAGGCCATCTCTCATCTTAAGTACT	1020
Qy	5125	CAGCACTAGCTATCTTGTGATGATCTTTTGTGTTGGTGGGGTGTACATTTGCTTGTTCAT	5184
Db	1021	CAGAACTAGCTATCTTGTACATCTCTCTGTCGTTGGGGTGTACATTTGCTTGTTCAT	1080
Qy	5185	GCGAGTTATTGCTGAGGAAGATCAAACTATGATCAAAAGGCTAAATTTTGAGGC	5244
Db	1081	GCGAGTCATCGCTGAGGAAGATCAAACTATGATCAAAAGGCTAAATTTTGAGGC	1140
Qy	5245	TCAAAGGAAAGGTATGGTTATAAACTATCTTTTGAATCTTTTAAAGATCTTCAAAGTG	5304
Db	1141	TCAAAGGAAAGGTATGGTCTATAACACTATCTCTTGTATCTCCCAAAGATCTTCAAAGTG	1200
Qy	5305	TGAGTATGTTTATTTGGTTGGCTTCTGGTGATATTATGTTTATAGAAATTTGCTCTTAT	5364
Db	1201	TGAGTATGTTTATTCGTTGGCTTCTGGTGATATTATGTTTATAGAAATTTGCTCTTAT	1260
Qy	5365	ATATTGGCTATATAGAGGTGCGGTGATATGATGAATCAAGAGTTGATTTGGAAA	5424
Db	1261	ATATTGGCTATATAGAGGTGCGGTGATATGATGAATCAAGAGTTGATTTGGAAA	1320
RESULT 4			
CW802337			
LOCUS	WiscdsLox477-480J24 Arabidopsis thaliana T-DNA insertion flanking		
DEFINITION	sequences Arabidopsis thaliana genomic, genomic survey sequence.		
ACCESSION	CW802337		
VERSION	CW802337.1 GI:56000165		
KEYWORDS	GSS.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 736)		
AUTHORS	Woody,S.T., Monson,S.S., Jester,P.J., Austin-Phillips,S., Anasino,R.M., Sussman,M.R. and Krysan,P.J.		
TITLE	A New Community Resource for Knocking-Out Small Genes and Tandemly-Duplicated Gene Families and for Mosaic Analysis in Arabidopsis		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Woody St Biotechnology Center University of Wisconsin-Madison 425 Henry Mall, Madison, WI 53706, USA Tel: (608) 262-4640 Email: swoody@facstaff.wisc.edu Class: TAIL-PCR.		
FEATURES			
source			
1..736			
/organism="Arabidopsis thaliana"			
/mol_type="genomic DNA"			
/cultivar="Col-O ecotype"			
/db_xref="taxon:3702"			
/tissue_type="seeds produced by primary (Basta-resistant) transformants"			
/clone_lib="Arabidopsis thaliana T-DNA insertion flanking sequences"			
/note="Vector: pDS-Lox; Sequence generated in the course of an Arabidopsis T-DNA tagging program. TAIL-PCR was used to generate sequencing templates that represent A.t. genomic DNA flanking the left border of the pDS-Lox T-DNA insert. PCR products were sequenced directly by using the			

REFERENCE	1 (bases 1 to 720)	Woody, S.T., Monson, S.S., Jester, P.J., Austin-Phillips, S., Amasino, R.M., Sussman, M.R. and Krysan, P.J. A New Community Resource for Knocking-Out Small Genes and Tandemly-Duplicated Gene Families and for Mosaic Analysis in Arabidopsis Unpublished (2004) Contact: Woody ST Biotechnology Center University of Wisconsin-Madison 425 Henry Mall, Madison, WI 53706, USA Tel: (608) 262-4640 Email: swoody@facstaff.wisc.edu Class: TAIL-PCR.
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES	Location/Qualifiers source 1..720 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /cultivar="Col-O ecotype" /db_xref="taxon:3702" /tissue_type="seeds produced by primary (Basta-resistant) transformants" /clone_lib="Arabidopsis thaliana T-DNA insertion flanking sequences" /note="Vector: pDS-Lox; Sequence generated in the course of an Arabidopsis T-DNA tagging program. TAIL-PCR was used to generate sequencing templates that represent A.t. genomic DNA flanking the left border of the pDS-Lox T-DNA insert. PCR products were sequenced directly by using the p745 primer 5' AACGTCGCAATGCTATTAAAGTTGTC 3"	
ORIGIN		
Query Match	9.4%;	Score 711; DB 10; Length 720;
Best Local Similarity	100.0%;	Pred. No. 4.3e-116;
Matches	711; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	4676	AGTGGTATGAGTTCTAGTGGTGGGATCTTGGATCGATCGAGAGATACCTCCATCAACAA 4735
Db	10	AGTGGTATGAGTTCTAGTGGTGGGATCTTGGATCGATCGAGAGATACCTCCATCAACAA 69
QY	4736	GCCTCAGCAATTCCTTCCTGCATCAACACTACCGGATTTGGTGCATCTTCAACGCGTTA 4795
Db	70	GCTCAGCAATTCCTTCCTGCATCAACACTACCGGATTTGGTGCATCTTCAACGCGTTA 129
QY	4796	TATCCATTACTAGAGGTAAAGGAGGTGTTAATCAAGGTGATTTCTCAACGAGAGTAGT 4855
Db	130	TATCCATTACTAGAGGTAAAGGAGGTGTTAATCAAGGTGATTTCTCAACGAGAGTAGT 189
QY	4856	GATTATCCAATCAGCTAATGTTTAAAGCCCTTGATGATTTTCTTCAAGCGGGTTAGC 4915
Db	190	GATTATCCAATCAGCTAATGTTTAAAGCCCTTGATGATTTTCTTCAAGCGGGTTAGC 249
QY	4916	GCACGCAAAACAGAAATGTGAAGCGGAGAGAGATCATCAGGATCGGGTAGGAGTGGG 4975
Db	250	GCACGCAAAACAGAAATGTGAAGCGGAGAGAGATCATCAGGATCGGGTAGGAGTGGG 309
QY	4976	GATGAGTGAATTAACCTTATCAAGAAACCTTTTGGGTAAATCAACATAAACTCAGGCAGG 5035
Db	310	GATGAGTGAATTAACCTTATCAAGAAACCTTTTGGGTAAATCAACATAAACTCAGGCAGG 369
QY	5036	AACGAGGAATACATCATATGGGAGGTAAACAGTTCTTTGGACCGGTTTCACTCCAAAC 5095
Db	370	AACGAGGAATACATCATATGGGAGGTAAACAGTTCTTTGGACCGGTTTCACTCCAAAC 429
QY	5096	TCACAGGCGCATCTCTCATTTCTAGTACTCAGCACTAGCTATTTCTTGATCATTTTGT 5155
Db	430	TCACAGGCGCATCTCTCATTTCTAGTACTCAGCACTAGCTATTTCTTGATCATTTTGT 489
QY	5156	TGGTTGGGGTGTACATTTGGTCTTGTCTATGCGAGTATTGCTGAGGAAGATCAAAACCATG 5215
Db	490	TGGTTGGGGTGTACATTTGGTCTTGTCTATGCGAGTATTGCTGAGGAAGATCAAAACCATG 549
QY	5216	CAGCTATATCCAAAGGCTTAATTTTGGAGCTCAAGAGAAAGATATGGTTATATAAATCTATCT 5275

ORIGIN	p745 primer 5' AACGTCGCAATGCTATTAAAGTTGTC 3''	
Query Match	9.5%;	Score 720.4; DB 10; Length 736;
Best Local Similarity	98.4%;	Pred. No. 9.2e-118;
Matches	724; Conservative	0; Mismatches 12; Indels 0; Gaps 0;
QY	518	TTATATATTAACCTTTCAGCCCTTCGTGAAAAATGCGCATGTTCTTTGGTGGAACTAG 577
Db	1	TTNANNNTNNAACCTTTCGCCCTTCGTGAAAAATGCGCATGTTCTTTGTTGGAACTAG 60
QY	578	GAATCTTCTTCTACTAAAGATTGGCATGCAAGTGTGAAGTTTCCATTGTTATCTACTCT 637
Db	61	GAATCTTCTTCTACTAAAGATTGGCATGCAAGTGTGAAGTTTCCATTGTTATCTACTCT 120
QY	638	CGATATTTTCCCAACCTTAAATATCTTGAAATTAAGATCAAAATTATATGAGAACATATAT 697
Db	121	CGATATTTTCCCAACCTTAAATATCTTGAAATTAAGATCAAAATTATATGAGAACATATAT 180
QY	698	TGATCATTTGTTTGAATAGCCCTTATATAGTGTGTAGATGTGAACCTATGGATACAAAAA 757
Db	181	TGATCATTTGTTTGAATAGCCCTTATATAGTGTGTAGATGTGAACCTATGGATACAAAAA 240
QY	758	CAATAGGAAGAAAAGCTGCGAGAGAGTGTGCTTTTCAACTATGATATGGGTCAT 817
Db	241	CAATAGGAAGAAAAGCTGCGAGAGAGTGTGCTTTTCAACTATGATATGGGTCAT 300
QY	818	GAGATGTACACTACAGATAAGCGAAGATTCCTGTTGTCATGAAATGTCTTACTAATAAAA 877
Db	301	GAGATGTACACTACAGATAAGCGAAGATTCCTGTTGTCATGAAATGTCTTACTAATAAAA 360
QY	878	AAAAACATGCACATTTCTATAAAGACGAATTTCTTTTAAATATAAATTTCTATATAAA 937
Db	361	AAAAACATGCACATTTCTATAAAGACGAATTTCTTTTAAATATAAATTTCTATATAAA 420
QY	938	ATAAAGATAAGTGCTCCTTTAAACACATGCAAAAGAAATATAGATTTTACCGTATCAGAT 997
Db	421	ATAAAGATAAGTGCTCCTTTAAACACATGCAAAAGAAATATATAGATTTTACCGTATCAGAT 480
QY	998	TTTTCATACAAATTTTATATTTTTCAGCTTTGAAAGATTAACATGACAACTGTATCGTGT 1057
Db	481	TTTTCATACAAATTTTATATTTTTCAGCTTTGAAAGATTAACATGACAACTGTATCGTGT 540
QY	1058	GTCTCTGTCTATTACCCCTAGAGAAGTGAACACTTTATGTATTGTCATACGG 1117
Db	541	GTCTCTGTCTATTACCCCTAGAGAAGTGAACACTTTATGTATTGTCATACGG 600
QY	1118	CGAGCTAGCTTCTTCCCTACTTGTGCAATAGATGAAGACATATCACTCAGGTTACGCTA 1177
Db	601	CGAGCTAGCTTCTTCCCTACTTGTGCAATAGATGAAGACATATCACTCAGGTTACGCTA 660
QY	1178	CTTCGAAGCGCAACATATCGACAAAAATTCGTTTTCAGCTCTATCATCTGCTTTTGAAGA 1237
Db	661	CTTCGAAGCGCAACATATCGACAAAAATTCGTTTTCAGCTCTATCATCTGCTTTTGAAGA 720
QY	1238	AAATATCAACATATC 1253
Db	721	AAATATCAACATATC 736

RESULT 5	CW797151	720 bp	DNA	linear	GSS 23-NOV-2004
LOCUS	WisCDeLox377-380K7 Arabidopsis thaliana T-DNA insertion flanking				
DEFINITION	sequences Arabidopsis thaliana genomic, genomic survey sequence.				
ACCESSION	CW797151				
VERSION	CW797151.1				
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				

RESULT 5	CW797151	720 bp	DNA	linear	GSS 23-NOV-2004
LOCUS	CW797151				
DEFINITION	WtaCDeLox377-380K7 Arabidopsis thaliana T-DNA insertion flanking sequences Arabidopsis thaliana genomic, genomic survey sequence.				
ACCESSION	CW797151				
VERSION	CW797151.1	GI:55994979			
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				





Db 133 |GATGTGACTCAACCAATACTAAGTTCTGTTACTTCAATAATATAGCCTTACTCAACCTC 192  
QY 4353 GCCATTTCGCAAAACATGTCGTGCTATTGGACACGTCGGGTTCCTTGAGGAATGTC 4412  
Db 193 GCCATTTCGCAAAACATGTCGTGCTATTGGACACGTCGGGTTCCTTGAGGAATGTC 252  
QY 4413 CTGTTGGAGGAGGCTTTAGGAGGAACAGAGAAAGCAAAATCCAGATCGAAATCTACGGTCG 4472  
Db 253 CTGTTGGAGGAGGCTTTAGGAGGAACAGAGAAAGCAAAATCCAGATCGAAATCTACGGTCG 312  
QY 4473 TGGTCTCGACTGATAATACTACTAGTACTTCACTACTTCTCTCGCCCAAGTTACTCAA 4532  
Db 313 TGGTCTCGACTGATAATACTACTAGTACTTCACTACTTCTCTCGCCCAAGTTACTCAA 372  
QY 4533 ACCCTAGCAAGTTTCATAGCTACGCTCAATCCCGGAGTTTAATTCCAACTGCCCCATCT 4592  
Db 373 ACCCTAGCAAGTTTCATAGCTACGCTCAATCCCGGAGTTTAATTCCAACTGCCCCATCT 432  
QY 4593 TGCCTCCTCTCCAAAGCCTTGGAGATTACAATTCAAGCAACACTCGATTAGATTGTTGGTG 4652  
Db 433 TGCCTCCTCTCCAAAGCCTTGGAGATTACAATTCAAGCAACACTCGATTAGATTGTTGGTG 492  
QY 4653 GAACTCAAAATAGCAACATGATAAGTGGTATGAGTTCTAGTGGTGGGATCTTGGATGCAT 4712  
Db 493 GAACTCAAAATAGCAACATGATAAGTGGTATGAGTTCTAGTGGTGGGATCTTGGATGCAT 552  
QY 4713 GGAGAAATACCTCATCAACAAGCTCAGCAATTCCTTCTTGATCAACACTACCGGAT 4772  
Db 553 GGAGAAATACCTCATCAACAAGCTCAGCAATTCCTTCTTGATCAACACTACCGGAT 612  
QY 4773 TGGTGCATCTTCAACCGGTTATATCCATTACTAGAGGTAAGGAGGTGTTAATCAAG 4832  
Db 613 TGGTGCAATCTTCAACCGGTTATATCCATTACTAGAGGTAAGGAGGTGTTAATCAAG 649  
QY 4833 GTGATTTCTCAACAGAAGATAGTGTATTTCAATCAGCTAATGTTTAAGCCCTTGTATGG 4892  
Db 650 ----- 649  
QY 4893 ATTTTCTTCAGCGGGTTAGCGCCAGCAAAAGAAATGTGAAGCGGAGGAATG 4952  
Db 650 -----AGCGGGGTTAGCGCCAGCAGCAGCAAGAAATGTGAAGCGGAGGAATG 699  
QY 4953 ATCAGATCGGGTAGGAT-GGGATGGAGTGAATACTTATCAAGAAACTTTTGGGT 5011  
Db 700 ATCAGATCGGGTAGGATGGGATGGAGTGAATACTTATCAAGAAACTTTTGGGT 759  
QY 5012 AATATCAACATAAACTCAGCGCAGGAACGAGGAATACATCATCTGGGAGGTAAACAGTTCT 5071  
Db 760 AATATCAACATAAACTCAGCGCAGGAACGAGGAATACATCATCTGGGAGGTAAACAGTTCT 819  
QY 5072 TGGACCGGTTTCACTCCCAACAACTCAACAGGCCATCTCTCATTTCTA 5118  
Db 820 TGGACCGGTTTCACTCCCAACAACTCAACAGGCCATCTCTCATTTCTA 866

RESULT 8  
DR750041  
LOCUS  
DEFINITION  
CDNA clone 006-G10-Sela MP1Z-ADIS-065d Arabidopsis thaliana  
DR750041  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE  
AUTHORS  
1 (bases 1 to 866)  
Paz-Ares, J., Valencia, A., Costantino, P., Vittorioso, P., Davies, B.,  
Gilmartin, P., Giraudat, J., Parcy, F., Reindl, A., Sablowski, R.,  
Coupland, G., Martin, C., Angenent, G.C., Baeumlein, H., Mock, H.P.,

TITLE  
JOURNAL  
COMMENT

Carbonero, P., Colombo, L., Tonelli, C., Engstrom, P.,  
Droegge-Laser, W., Gatz, C., Kavanagh, T., Kushnir, S., Zabeau, M.,  
Laux, T., Holdsworth, M., Ruberti, I., Ratcliff, F., Smeekens, S.,  
Somssich, I., Weisshaar, B. and Traas, J.  
REGIA, an EU project on functional genomics of transcription  
factors from Arabidopsis thaliana  
Comp. Funct. Genomics 3 (2), 102-108 (2002)  
Contact:  
Paz-Ares, Costantino, Vittorioso, Davies, Gilmartin, Giraudat, Parcy, Sablowski,  
Coupland, Martin, Angenent, Baeumlein, Carbonero, Colombo, Tonelli,  
Engstrom, Droegge-Laser, Gatz, Kavanagh, Kushnir, Zabeau, Laux, Holdsworth,  
Ruberti, Smeekens, Somssich, Weisshaar, Traas  
Bielefeld University, Institute for Genome Research  
Universitaetsstrasse 25, D-33594 Bielefeld, Germany  
Email: bernd.weisshaar@uni-bielefeld.de  
AGI: AT3955370; SeqAnalysis: truncated in 5' at pos 179;  
Translation: no full cds detected  
Data analysis performed in the frame of REGULATORS (Exploiting  
inter-species conservation in promoter sequences to identify  
regulators of reproductive development and physiological  
performance), a Trilateral Co-Operation in Plant Genomics between  
Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by  
G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent  
Thareau (IBP-Orsay UMR8618 CNRS-UPS, thareau-ad-ibp.u-psud.fr) and  
Alain Lecharny (URGV-Evry UMR INRA-CNRS-UEVE,  
lecharny-ad-ibp.u-psud.fr).  
Definition of the terms used to describe the quality of the clone:  
The about 2250 sequences from the clone collection were sorted  
according to clones and clustered. If more than one contig was  
formed, the clone was designated 'contamination'. The contigs and  
singletons were blasted against CDS plus pseudogenes from the  
TIGRv5 annotation, and the resulting AGI code is presented if more  
than 90 percent identity was found. The sequences were also blasted  
against all TIGRv5 introns, and matches longer than 50 bp with 95  
percent identity are reported as 'intron found'. The remaining  
terms for SeqAnalysis describe the outcome of the evaluation of the  
CDS detected after pairwise alignment with CDS plus pseudogenes  
from the TIGRv5 annotation file. The sequences or contigs for which  
a full CDS with or without STOP codon was detected, a BLASTp  
against all TIGRv5 protein sequences was performed. Full perfect:  
100 percent identity; full good: better than 95 percent identity  
over more than 95 percent of the sequence; partial good: better  
than 95 percent identity over less than 95 percent of the sequence;  
weak similarity: less than 95 percent identity over less than 95  
percent of the sequence; no similarity: no hit from BLASTp. Note  
that the collection contains a few clones for which sequencing was  
not successful, which could have simple technical reasons.  
Obviously, information about these clones is missing in the  
submitted data.

Insert Length: 866 Std Error: 0.00  
Seq primer: Sela TCGCGTTAAGCTAGCATGATCTC.

FEATURES  
source

Location/Qualifiers  
1..866  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="GABI:1452132"  
/db\_xref="taxon:3702"  
/clone="006-G10"  
/lab\_host="E. coli DH5alpha"  
/clone\_lib="MP1Z-ADIS-065d"  
/note="Vector: pDONR201; In the context of the EU-funded  
project REGIA (QLG-CT11999-00876, coordinator Javier  
Paz-Ares), a set of transcription factor ORFs was  
generated. The ORFs were produced in a decentralized way  
in the labs of the participants. Most of the ORFs were  
generated by RT-PCR using cDNA from various A. thaliana  
tissues as a template. Initially, it was planned to use  
yeast recombination to move the ORFs from the cloning  
vectors into target constructs. For this reason, a number  
of the clones contain 'REGIA tags' (RG\_tag1:  
5pr-AATTCCAGCTCACCACC-3pr; RG\_tag2:  
5pr-CATGGCAATTCCCGGATC-3pr). During the lifetime of the



project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequencing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weisshaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by GabiPD (<http://gabi.rzpd.de>)."

```
ORIGIN
Query Match      8.9%; Score 671.4; DB 8; Length 866;
Best Local Similarity 88.1%; Pred. No. 4.8e-109;
Matches 781; Conservative 0; Mismatches 12; Indels 93; Gaps 1;

QY 4233 TGGTGGAGCTGCTCGGATCGCAAAAGTCCCAATGCTGCTGAAGCAGCTCTAAATGCCCTA 4292
DB 18 TGGNGGACCTTGAAGAAGATCGCAAAAGTCCCAATGCTGCTGAAGCAGCTCTAAATGCCCTA 77

QY 4293 GATGTGACTCAACCAATAGTCTGTGTTACTTCAATAACTATAGCCTTACTCAACCTC 4352
DB 78 GATGTGACTCAACCAATAGTCTGTGTTACTTCAATAACTATAGCCTTACTCAACCTC 137

QY 4353 GCCATTTCTGAAAAATGTCGTCTATTTGACACGTCGCGGTTCTTTGAGGAATGTTT 4412
DB 138 GCCATTTCTGAAAAATGTCGTCTATTTGACACGTCGCGGTTCTTTGAGGAATGTTT 197

QY 4413 CTGTTGGAGGAGCTTTAGGAGGACAGAGAGCAATCCAGATCGAATCTACGTCG 4472
DB 198 CTGTTGGAGGAGCTTTAGGAGGACAGAGAGCAATCCAGATCGAATCTACGTCG 257

QY 4473 TGGTCTCGACTGATAATCTAGTACTTCTATCTACTTCTCGCCCAAGTTACTCAA 4532
DB 258 TGGTCTCGACTGATAATCTAGTACTTCTATCTACTTCTCGCCCAAGTTACTCAA 317

QY 4533 ACCCTAGCAAGTTTCATAGCTACGTCGCAATCCCGAGTTTAATTCACATGCCCCT 4592
DB 318 ATTTTAGCAAGTTTCATAGCTACGTCGCAATCCCGAGTTTAATTCACATGCCCCT 377

QY 4593 TGCCTCTCTCAAGACCTTGAGATTAACAAATCAAGCAACTGGATAGATTTGGTG 4652
DB 378 TGCCTCTCTCAAGACCTTGAGATTAACAAATCAAGCAACTGGATAGATTTGGTG 437

QY 4653 GAACTCAATAAGCAACATGATAGTGGTATGATGATCTAGTGGGATCTTGGATGAT 4712
DB 438 GAACTCAATAAGCAACATGATAGTGGTATGATGATCTAGTGGGATCTTGGATGAT 497

QY 4713 GGAGAAATACCTCCATCACAAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGAT 4772
DB 498 GGAGAAATACCTCCATCACAAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGAT 557

QY 4773 TGGTGAATCTTCAACCGGTTATATCCATTAAGAGGTGAAGAGGTGTTAATCAAG 4832
DB 558 TGGTGAATCTTCAACCGGTTATATCCATTAAGAGGTGAAGAGGTGTTAATCAAG 594

QY 4833 GTGATTTCTCAACAGAGAGTAGTGTATTTTCAATCAGCTAATGTTTAAGCCCTTGATGG 4892
DB 595 ----- 594

QY 4893 ATTTTCTTTTTCAGCGGGGTTAGCGCCAGCAAAAGAAATGTGAAGCGGGAAGAAATG 4952
DB 595 -----AGCGGGGTTAGCGCCAGCAAAAGAAATGTGAAGCGGGAAGAAATG 644

QY 4953 ATCAGATCGGGGTAGGATGGGATGGAGTGAATTAATCAAGAACTTTTGGGTA 5012
DB 645 ATCAGATCGGGGTAGGATGGGATGGAGTGAATTAATCAAGAACTTTTGGGTA 704

QY 5013 ATATCAACATAACTCAGGAGGAGGAGGATACATCATGCGGAGGTAACTGTTT 5072
DB 705 ATATCAACATAACTCAGGAGGAGGAGGATACATCATGCGGAGGTAACTGTTT 764
```

```
QY 5073 GGACGGTTTTCACCTCCCAACCACTCAACAGGCCATCTCTCATTTCTA 5118
DB 765 GGACGGTTTTCACCTCCCAACCACTCAACAGGCCATCTCTCATTTCTA 810

RESULT 9
LOCUS CX308688/c 733 bp mRNA linear EST 06-MAY-2005
DEFINITION CX308688 Citrus clementina cDNA clone C21002G02, mRNA
sequence.
ACCESSION CX308688
VERSION CX308688.1 GI:63077542
KEYWORDS EST.
SOURCE Citrus clementina
ORGANISM Citrus clementina
REFERENCE 1 (bases 1 to 733)
AUTHORS Forment,J., Gadea,J., Huerta,L., Abizanda,L., Agustí,J., Alamar,S.,
Alos,E., Andres,F., Arribas,R., Beltran,J.P., Berbel,A.,
Blazquez,M.A., Brumos,J., Canas,L.A., Cercos,M.,
Colmenero-Flores,J.M., Conesa,A., Establies,B., Gandia,M.,
Garcia-Martinez,J.L., Gimeno,J., Gisbert,A., Gomez,G.,
Gonzalez-Candelas,L., Granel,A., Guerri,J., Lafuente,M.T.,
Madueno,F., Marcos,J.F., Marques,M.C., Martinez,F.,
Martinez-Godoy,M.A., Miralles,S., Moreno,P., Navarro,L., Pallas,V.,
Perez-Amador,M.A., Perez-Valle,J., Pons,C., Rodrigo,I.,
Rodriguez,P.L., Royo,C., Serrano,R., Soler,G., Tadeo,F., Talon,M.,
Terol,J., Trenor,M., Vaello,L., Vicente,O., Vidal,Ch., Zacarias,L.
and Conejero,V.
TITLE Development of a citrus genome-wide EST collection and cDNA
microarray as resources for genomic studies
JOURNAL Plant Mol. Biol. 57 (3), 375-391 (2005)
PUBMED 15830128
COMMENT Contact: Forment J
Genomics Laboratory
Instituto de Biologia Molecular y Celular de Plantas (Universidad
Politcnica de Valencia - Consejo Superior de Investigaciones
Cientificas)
Avenida de los Naranjos s/n, 46022 Valencia, Spain
Email: jforment@bmcp.upv.es.
FEATURES
source
location/Qualifiers
1..733
/organism="Citrus clementina"
/mol_type="mRNA"
/cultivar="Clemenules"
/db_xref="taxon:85681"
/clone="C21002G02"
/sex="hermaphrodite"
/dev_stage="adult trees"
/lab_host="Escherichia coli"
/clone_lib="AbeLeaSub1"
/organ="leaves; Vector: pCR2.1; Subtracted cDNA
library made from laminar abscission zone of leaves"
ORIGIN
Query Match 7.5%; Score 570.2; DB 8; Length 733;
Best Local Similarity 92.5%; Pred. No. 5e-91;
Matches 615; Conservative 0; Mismatches 38; Indels 12; Gaps 1;

QY 6190 AGATATCATCATCAATCCACTTGTGTTGAAGACGTGGTTGGAACTGCTTTTTCACGA 6249
DB 679 AGATATCATCATCAATCCACTTGTGTTGAAGACGTGGTTGGAACTGCTTTTTCACGA 620

QY 6250 TGTCTCTGTTGGTGGGGTCCATCTTTGGACCACTGTCGGTAGAGCATCTTGAACGA 6309
DB 619 TGTCTCTGTTGGTGGGGTCCATCTTTGGACCACTGTCGGTAGAGCATCTTGAACGA 560

QY 6310 TAGCCTTCTTTTATCGCAATGATGGCATTTGTAGAAGCATCTTCTTTTCTACTGTCC 6369
DB 559 TGGCCTTCTTTTATCGCAATGATGGCATTTGTAGAAGCATCTTCTTTTCTACTGTCT 500
```





/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="1119 - RescueMu Grid AA"  
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.mdb.iastate.edu' and follow the links for 'RescueMu.' Grid AA was grown at UC San Diego in 2002. DNA was extracted from leaf strips, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

ORIGIN

Query Match 6.3%; Score 477.4; DB 10; Length 597;  
Best Local Similarity 90.8%; Pred. No. 1.6e-74;  
Matches 524; Conservative 0; Mismatches 41; Indels 12; Gaps 1;  
QY 6284 ACTGTCGGTAGGCGATCTTGAACGATAGCCCTTCTTCATCGCAATGATGGCATTTGTA 6343  
Db ACTGTCGGTAGGCGATCTGCAACGATGSCCTTCTTCATCGCAATGATGGCATTTGTA 538  
QY 6344 GAAGCGATCTCTCTTTCTACTGTCCTTTTCGATGAGTGACAGATAGCTGGGCAATGGAA 6403  
Db GGAGCGCTCTGCTTTTCCACTATCTTCAACAATAAGTGACAGATAGCTGGGCAATGGAA 478  
QY 6404 TCCGAGGAGTTTCCCGATATTACCTTTGTTGAAAGTCTCAATAGCCCTCTGCTCTTC 6463  
Db TCCGAGGAGTTTCCCGATATTACCTTTGTTGAAAGTCTCAATAGCCCTTTGCTCTTC 418  
QY 6464 TGAGACTGTATCTTTGATATCTTGGAGTAGACGAGTGTCTGCTCCACCATGTGGG 6523  
Db TGAGACTGTATCTTTGATATCTTGGAGTAGACGAGTGTCTGCTCCACCATGT --- 362  
QY 6524 GATCTAGATATCATATCAATCACTTGTCTTGAAGACGTTGTTGAAACGCTCTTTTTC 6583  
Db TATCATATCAATCACTTGTCTTGAAGACGTTGTTGAAACGCTCTTTTTC 310  
QY 6584 CACGATGTTCTCGTGGGGTCCATCTTTGGGACCACTGTCGTAGAGGCATCTTG 6643  
Db CACGATGTTCTCGTGGGGTCCATCTTTGGGACCACTGTCGTAGAGGCATCTTC 250  
QY 6644 AACGATAGCTTTCTCTTTATCGCAATGATGGCATTTGTAGAGCCATCTTCTTTTCTAC 6703  
Db AACGATAGCTTTCTCTTTATCGCAATGATGGCATTTGTAGAGCCATCTTCTTTTCCAC 190  
QY 6704 TGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATAT 6763  
Db TATCTTCAATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATAT 130  
QY 6764 TACCTTTTGAAGTCTCAATAGCCCTCTGCTTCTGAGACTGTATCTTTGATATT 6823  
Db TACCTTTTGAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATT 70  
QY 6824 CTTGAGTAGACGAGAGTGTCTGCTCCACCATGTTG 6860  
Db TTTGAGTAGACAGTGTCTGCTCCACCATGTTG 33

RESULT 12  
B26529  
LOCUS B26529 486 bp DNA linear GSS 13-OCT-1997  
DEFINITION F9G12TF IGF Arabidopsis thaliana genomic clone F9G12, genomic survey sequence.  
ACCESSION B26529  
VERSION B26529.1 GI:2512495  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 486)  
Rounsley,S.D., Kelley,J.M., Field,C.E., Craven,M.B., Adams,M.D. and Venter,J.C.  
Use of a BAC End Sequence Database To Identify Minimal Overlaps for Arabidopsis Genomic Sequencing  
Unpublished (1997)  
Contact: Steve Rounsley  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: rounsley@tigr.org  
Class: BAC ends  
Seq primer: M13-21  
High quality sequence stop: 486.  
Location/Qualifiers  
1..486  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="F9G12"  
/sex="hermaphrodite"  
/clone\_lib="IGF"  
/note="Vector: BelobACII; Site\_1: EcoRI; Site\_2: EcoRI;  
Produced by Thomas Altman"

ORIGIN

Query Match 6.1%; Score 461.2; DB 9; Length 486;  
Best Local Similarity 98.1%; Pred. No. 1.3e-71;  
Matches 477; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
QY 2024 ATTTTCTTAAACAGAGGAATTTTACATCATCTTTAGACTGAACCTTCGGGATAAAAA 2083  
Db ATTTTGGTCTTAAACAGAGGAATTTTACATCATCTTTAGACTGAACCTTCGGGATAAAAA 60  
QY 2084 TCTGCCATGCAAAAGGTAATTTATTTTTCATGACAAAAGCCACAAATGCGGATAATATA 2143  
Db TCTGCCATGCAAAAGGTAATTTATTTTTCATGACAAAAGCCACAAATGCGGATAATATA 120  
QY 2144 ACTATAATATATGCAAAACGAAACTTTTACCTTGGGTCTATACCGAGGAAAAAAGGTACA 2203  
Db ACTATAATATATGCAAAACGAAACTTTTACCTTGGGTCTATACCGAGGAAAAAAGGTACA 180  
QY 2204 CTGGATTGTGCAACTCCACCAAGACCAACCCACTACCAATTCACCTTTATTTCTTCT 2263  
Db CTGGATTGTGCAACTCCACCAAGACCAACCCACTACCAATTCACCTTTATTTCTTCT 240  
QY 2264 TTATTTCACTCAAAATCTTTTAAATTTTAAATTAATTAATTAATTCCTCTCTGTTG 2323  
Db TTATTTCACTCAAAATCTTTTAAATTTTAAATTAATTAATTAATTCCTCTCTGTTG 300  
QY 2324 TTTTAAACCTTTTAAATTAATAATGAAAGAGGTGCTAGGGTTCTCTCTCGATGG 2383  
Db TTTTAAACCTTTTAAATTAATAATGAAAGAGGTGCTAGGGTTCTCTCTCGATGG 360  
QY 2384 CCACCTCTGCTCTTCAATCTTTTGGGACCACTTTGACGATGCGAAGAAAAAAA 2443  
Db CCACCTCTGCTCTTCAATCTTTTGGGACCACTTTGACGATGCGAAGAAAAAAA 420  
QY 2444 AAAAAAATAAAACCTT-ACATGCTAAGAAAACTCTCTTCTTGTGTCTCTCTCATATA 2502  
Db GTAAAAAAGTAAACCTTAACTGCTAAGAAAACTCTCTTCTTGTGTCTCTCATATA 480  
QY 2503 AAAGAA 2508  
Db AAAGAA 486

AQ011657 452 bp DNA linear GSS 29-MAY-1998  
LOCUS F24E16TFB IGF Arabidopsis thaliana genomic clone F24E16, genomic  
DEFINITION survey sequence.  
ACCESSION AQ011657  
VERSION AQ011657.1 GI:3166407  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 452)  
AUTHORS Rounsley, S.D., Suh, E.J., Wible, C., Golden, K., Shatsman, S., Choi, P.,  
Yu, K., Akinretaye, B., Shen, K., Goonasekaram, S., Militscher, J.,  
Adams, M.D. and Venter, J.C.  
TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in  
Arabidopsis Genomic Sequencing. Update 4  
JOURNAL Unpublished (1998)  
COMMENT Other\_GSSs: F24E16TR  
Contact: Steve Rounsley  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: rounsley@tigr.org  
Seq primer: M13-21  
Class: BAC ends  
High quality sequence stop: 452.  
Location/Qualifiers  
1. .452  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="F24E16"  
/sex="hermaphrodite"  
/clone\_lib="IGF"  
/note="Vector: BelobACII; Site\_1: EcoRI; Site\_2: EcoRI;  
Produced by Thomas Altmann"  
ORIGIN  
Query Match 5.8%; Score 442; DB 9; Length 452;  
Best Local Similarity 98.9%; Pred. No. 3.3e-68;  
Matches 445; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 4231 AATGGTGGAACTGCTCGGATCGCAAAAGTCCCATTTGCTGAAGCAGCTCTAAATTTGCC 4290  
DB 3 AATGGTGGCAGCTGCTCGGATCGCAAAAGTCCCATTTGCTGAAGCAGCTCTAAATTTGCC 62  
QY 4291 TAGATGTGACTCAACCACTACTAAGTTCGTGTACTTCAATACTATAGCTTACTCAACC 4350  
DB 63 TAGATGTGACTCAACCACTACTAAGTTCGTGTACTTCAATACTATAGCTTACTCAACC 122  
QY 4351 TCGCCATTTCTCGAAACATGTCGTGCTATTGGACACGTGCGGTTCTTTCGAGAATGT 4410  
DB 123 TCGCCATTTCTCGAAACATGTCGTGCTATTGGACACGTGCGGTTCTTTCGAGAATGT 182  
QY 4411 TCCTGTTGGAGAGGCTTTAGGAGAACCAAGAGAGCAATCCAGATCGAAATCTACGGT 4470  
DB 183 TCCTGTTGGAGAGGCTTTAGGAGAACCAAGAGAGCAATCCAGATCGAAATCTACGGT 242  
QY 4471 CGTGGTCTCGATGATACTACTAGTACTTTCATCCTTACTTCTCGCCCAAGTTACTC 4530  
DB 243 CGTGGTCTCGATGATACTACTAGTACTTTCATCCTTACTTCTCGCCCAAGTTACTC 302  
QY 4531 AAACCCCTAGCAAGTTTCATAGCTACGGTCAAAATCCCGGAGTTTAATCCAACTTCCCAT 4590  
DB 303 AAACCCCTAGCAAGTCTCATAGCTACGGTCAAAATCCCGGAGTTTAATCCAACTTCCCAT 362  
QY 4591 CTTGCTCTCTCCAAAGCTTTGGAGATTACAACTTCAAGCAACACTGGATTAGATTTCG 4650

DB 363 CTTGCTCTCTCCAAAGCTTTGGAGATTACAACTTCAAGCAACACTGGATTAGATTTCG 422  
QY 4651 TGGAACTCAATAAGCAACATGATAAGTGG 4680  
DB 423 TGGAACTCATATACCGCAACATGATACGTGG 452  
RESULT 14  
B97814 433 bp DNA linear GSS 31-MAR-1998  
LOCUS F18P18TFB IGF Arabidopsis thaliana genomic clone F18P18, genomic  
DEFINITION survey sequence.  
ACCESSION B97814  
VERSION B97814.1 GI:2999893  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 433)  
AUTHORS Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Linher, K.,  
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Adams, M.D. and  
Venter, J.C.  
TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in  
Arabidopsis Genomic Sequencing. Update 3  
JOURNAL Unpublished (1997)  
COMMENT Other\_GSSs: F18P18TR  
Contact: Steve Rounsley  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: rounsley@tigr.org  
Seq primer: M13-21  
Class: BAC ends  
High quality sequence stop: 433.  
Location/Qualifiers  
1. .433  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="F18P18"  
/sex="hermaphrodite"  
/clone\_lib="IGF"  
/note="Vector: BelobACII; Site\_1: EcoRI; Site\_2: EcoRI;  
Produced by Thomas Altmann"  
ORIGIN  
Query Match 5.7%; Score 431.4; DB 9; Length 433;  
Best Local Similarity 99.8%; Pred. No. 2.5e-66;  
Matches 432; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 5459 CTCATTTCTTGGAGACCCATTATGAGACATTTGAGACATCTATAGCAATATATGTAATG 5518  
DB 1 CTCATTTCTTGGAGACCCATTATGAGACATTTGAGACATCTATAGCAATATATGTAATG 60  
QY 5519 TATATTAACCGTACTTAAAGTCGAATTTTATGACCAAGTAAATAAATATATGCCGAATGTA 5578  
DB 61 TATATTAACCGTACTTAAAGTCGAATTTTATGACCAAGTAAATAAATATATGCCGAATGTA 120  
QY 5579 CATGTAATATCGAGTTTAAACTATTTTTCGAATATATAACAACATTTTCTCTTCGTC 5638  
DB 121 CATGTAATATCGAGTTTAAACTATTTTTCGAATATATAACAACATTTTCTCTTCGTC 180  
QY 5639 AACTTATATACCTTATTTCTGATCTTATTTCTTTTAAATTCCTTTTCCTTCC 5698  
DB 181 AACTTATATACCTTATTTCTGATCTTATTTCTTTTAAATTCCTTTTCCTTCC 240  
QY 5699 AAGACACAAAAAATAAATACAGAAACGAAAAAGAGATTTTAAATAATTCATAACCC 5758

Db 241 AAGACACAAAAAATAACAGAAAGAAAAAGAGATTTTAAAAATTCATAACCC 300  
 QY 5759 ACGAGAAATATGACCTAAATTCAGACTAATCCCAAAATTCAGAAATTTATGTAATTT 5818  
 Db 301 ACGAGAAATATGACCTAAATTCAGACTAATCCCAAAATTCAGAAATTTATGTAATTT 360  
 QY 5819 TCGGATTTAATATGTTGTTCAAAATCAATGCGCAACTAACTAATTTGAAAGACAATGG 5878  
 Db 361 TCGGATTTAATATGTTGTTCAAAATCAATGCGCAACTAACTAATTTGAAAGACAATGG 420  
 QY 5879 AATGACTGAAACC 5891  
 Db 421 AATGACTGAAACC 433

## RESULT 15

CL517384  
 LOCUS SACS3B03 Flanking Sequence Tag of Oryza sativa T-DNA insertion lines  
 DEFINITION Oryza sativa (japonica cultivar-group) genomic, genomic survey  
 sequence.

ACCESSION CL517384  
 VERSION CL517384.1 GI:46144159

KEYWORDS Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzoae; Oryza.

REFERENCE 1 (bases 1 to 468)  
 AUTHORS Sallaud, C., Gay, C., Larmande, P., Bes, M., Piffanelli, P., Piegou, B.,  
 Droc, G., Regad, F., Bourgeois, E., Meynard, D., Perin, C.,  
 Ghesquiere, A., Delseny, M., Glaszmann, J.C. and Guiderdoni, E.  
 TITLE High throughput T-DNA insertion mutagenesis in rice: A first step  
 towards in silico reverse genetics  
 JOURNAL Plant J. (2004) In press  
 COMMENT Contact: Guiderdoni  
 UMR PIA Biotrop program  
 CIRAD

TA 40/03 ave Agropolis 34398 Montpellier cedex 5 FRANCE

Tel: 33467615629

Fax: 33467615605

Email: emmanuel.guiderdoni@cirad.fr

Class: TDNA tagged.

Location/Qualifiers

1. 468 /organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="Genomic DNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone\_lib="Flanking Sequence Tag of Oryza sativa T-DNA

insertion lines"

/note="PCR was performed on DNA of primary transformants  
 of Oryza sativa plants. The DNA fragment(s) resulting of  
 PCR were directly sequenced from the left border to  
 determine the genomic sequence flanking the insertion.  
 T-DNA derived sequences were removed. Information to order  
 the corresponding mutant line and a link to a database  
 providing a graphical display is available from June 2004  
 at <http://genoplante-info.infobiogen.fr/oryzatagline/>.  
 This sequence has been generated in the framework of the  
 French plant genomics program Genoplante  
 (<http://www.genoplante.org> and  
<http://genoplante-info.infobiogen.fr>)."

## ORIGIN

Query Match 5.6%; Score 424.8; DB 10; Length 468;  
 Best Local Similarity 94.6%; Pred. No. 3.8e-65;  
 Matches 456; Conservative 0; Mismatches 12; Indels 14; Gaps 1;

QY 6279 GGACCACTGTGGTAGAGGCATCTTGAACGATAGCGCTTTCCTTTATCGCAATGATGGCAT 6338  
 Db 1 GGACCACTGTGGCAGAGGCATCTTGAACGATAGCGCTTTCCTTTATCGCAATGATGGCAT 60

QY 6339 TTGTAGAAAGCCATCTTCTCTTTCTACTGTCTCTTTTCATGAAGTGAAGATAGCTGGGCAA 6398  
 Db 61 TTGTAGGTGCCACCTTCTCTTTCTACTGTCTCTTTTGTATGAAGTGAAGATAGCTGGGCAA 120  
 QY 6399 TGGAAATCCGAGAGGTTTCCCGATATTACCTTTTGTGAAAAGTCTCAATAGCCCTCTGG 6458  
 Db 121 TGGAAATCCGAGAGGTTTCCCGATATTACCTTTTGTGAAAAGTCTCAATAGCCCTCTGG 180  
 QY 6459 TCTTCTGAGACTGTATCTTTTGATATTTCTGGAGTAGACGAGAGTGTCTGCTCCACCATG 6518  
 Db 181 TCTTCTGAGACTGTATCTTTTGATATTTCTGGAGTAGACGAGAGTGTCTGCTCCACCATG 240  
 QY 6519 TTGGGGATCTAGATATCAATCAATCTCTTTTGAAGAGCTGGTTGGAAAGCTCTTCT 6578  
 Db 241 T-----TCACATCAATCCACTTGTCTTTGAAGAGCTGGTTGGAAAGCTCTTCT 286  
 QY 6579 TTTTCCACGATGTTCTCTGCTGGGGGTCCATCTTTTGGGACCACTCTCGTAGAGGCA 6638  
 Db 287 TTTTCCACGATGTTCTCTGCTGGGGGTCCATCTTTTGGGACCACTCTCGTAGAGGCA 346  
 QY 6639 TCTTGAACGATAGCTTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTCTTT 6698  
 Db 347 TCTTGAACGATAGCTTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTCTTT 406  
 QY 6699 TCTACTGTCTTTTCGATGAAGTGAAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCC 6758  
 Db 407 TCTACTGTCTTTTGTATGAAGTGAAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCC 466  
 QY 6759 GA 6760  
 Db 467 GA 468

Search completed: December 31, 2005, 05:09:55

Job time : 17838 secs

**BEST AVAILABLE COPY**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2005, 15:08:00 ; Search time 24337 Seconds  
(without alignments)  
17704.457 Million cell updates/sec

Title: US-10-650-249-1

Perfect score: 7580

Sequence: 1 agctctattaatcaagaga.....gcgcacacgggtgagct 7580

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_ev.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6131.2	80.9	103240	15	ATT22E16
c	2	1402.4	18.5	10138	11 AF187951
c	3	1402.4	18.5	10450	11 AF218466
4	1163	15.3	1457	15	AK221402
5	1156.2	15.3	1274	15	AF155818
c	6	640.2	8.4	9663	11 AF406991
c	7	640.2	8.4	12537	11 AY788908
c	8	640.2	8.4	12607	11 AY737283
c	9	621	8.2	5033	6 I28266
10	617.2	8.1	8626	11	CVU10488
11	617.2	8.1	8731	11	CVU10491
12	617.2	8.1	8742	11	AF234315
13	617.2	8.1	8894	11	AF234313
14	617.2	8.1	8909	11	CVU10487
15	617.2	8.1	8958	11	AF234296
16	617.2	8.1	9014	11	CVU10490
17	617.2	8.1	9110	11	AF234292
18	617.2	8.1	9148	11	AF294976

19	617.2	8.1	9170	11	AF294977	AF294977 Cloning v	
20	617.2	8.1	9182	11	AF294978	AF294978 Cloning v	
21	617.2	8.1	9937	11	AY560325	AY560325 Cloning v	
22	617.2	8.1	9946	11	AY560326	AY560326 Cloning v	
23	617.2	8.1	10549	11	AF234298	AF234298 Binary ve	
24	617.2	8.1	11633	11	AF234316	AF234316 Binary ve	
c	25	617.2	8.1	11667	6	AX840053	AX840053 Sequence
26	617.2	8.1	11785	11	AF234314	AF234314 Binary ve	
27	617.2	8.1	11846	11	AF354045	AF354045 Binary ve	
28	617.2	8.1	11849	11	AF234297	AF234297 Binary ve	
29	617.2	8.1	11881	11	AY739899	AY739899 Cloning v	
30	617.2	8.1	11921	11	AF354046	AF354046 Binary ve	
31	617.2	8.1	12001	11	AF234293	AF234293 Binary ve	
32	617.2	8.1	12072	11	AF294981	AF294981 Binary ve	
33	617.2	8.1	12361	11	AF234299	AF234299 Binary ve	
34	617.2	8.1	12361	11	AF234300	AF234300 Binary ve	
c	35	617.2	8.1	12614	6	AX052539	AX052539 Sequence
36	617.2	8.1	12942	11	AF294982	AF294982 Binary ve	
37	617.2	8.1	14203	11	AF294979	AF294979 Binary ve	
38	617.2	8.1	14230	11	AF294980	AF294980 Binary ve	
39	616.8	8.1	8626	11	CVU10459	U10459 Binary clon	
40	616.8	8.1	8731	11	CVU10463	U10463 Binary clon	
41	616.8	8.1	8909	11	CVU10458	U10458 Binary clon	
42	616.8	8.1	9014	11	CVU10462	U10462 Binary clon	
43	616.8	8.1	12537	11	AY788908	AY788908 N-termina	
44	616.8	8.1	12607	11	AY737283	AY737283 C-termina	
45	616.6	8.1	8860	11	AF234301	AF234301 Binary ve	

#### ALIGNMENTS

ATT22E16 103240 bp DNA linear PLN 16-APR-2005  
Arabidopsis thaliana DNA chromosome 3, BAC clone T22E16.

AL132975 GI:6434228

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..103240

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/db\_xref="taxon:3702"

/ecotype="Columbia"

1..8852

/note="overlap to BAC T26112 (EMBL:AL132954)"

992..3164

/gene="T22E16.10"

Join(992..1181,2134..3164)

CDS

	<pre>/gene="T22E16.10" /note="similarity to predicted proteins, Arabidopsis thaliana" /codon_start=1 /product="putative protein" /protein_id="CAB75893.1" /db_xref="GI:7076778" /db_xref="UniProt/TREMBL:Q9M2U3" /translation="MGPIKTIKKKGAERKQDVRNVLLAATAAATSASAAAAANNDDDD DSNGSQLDWDFGRRYGGSTDPKTFESVFKISRKTFDYICSLVKADFTAKPANFS DSSGSLNDRAVALARLGSGLSVIGETFGNQSTVSOITWRFVSEERAIHH LSWPKLDEIKSKFKISGLPNCAGAI01THVMNLPAVEPSNKMVLGDEKFNMTLQ AVVDPMELDVIACWPGSLNDVDVLKNSGFYKLVKGRKLNGLKPLSERTELREYI VGDGFPLLPWLITPYQKQPSLPTFEFNKHSSEATKAAQALSCLKDRWRLINGVMW MPDRNRLPRIFVCCLLHNI1IDMEDQTLDDQPLSQQDHMMYRQRCKLADEASVLR DELSQQLCKKSSA" /992..1181 /gene="T22E16.10" /number=1 /intron 1182..2133 /gene="T22E16.10" /number=1 /exon 2134..3164 /gene="T22E16.10" /number=2 /gene="T22E16.20" 3660..5330 /gene="T22E16.20" complement(join(3660..3728,3821..4575,5038..5077, 5262..5330)) /gene="T22E16.20" complement(join(3660..3728,3821..4575,5038..5077, 5262..5330)) /gene="T22E16.20" /note="strong similarity to synaptic glycoprotein Sc2 spliced variant, Homo sapiens, EMBL:AF038958" /codon_start=1 /product="glycoprotein-like" /protein_id="CAB75894.1" /db_xref="GI:7076779" /db_xref="GOA:Q9M2U2" /db_xref="InterPro:IPR001104" /db_xref="UniProt/TREMBL:Q9M2U2" /translation="MKVTVVSRSREVLKAPLDLPDSATVADLQAFHKKRKKFYPSR QRLTLPVTPGSKDPVNLNSKSKLKEYCDGNNSLTVFKDLGAQVSYLTLEFFVYLIG PLLIYFVYFVYFLYKYGEDCVIHPVQTYAMYIWCFFYKRILETFFVHRFSATS PTGNVFRNCAYTWSGAYIAYVNHPLVTPVSDLQMKIGFGGLVCQVAFYCHILLK NLRDPSGAGGYQIPRGFLFNIYVCANYTTEIYQWLGFENIATQTIAGYVFLAVALIMT NWLKGHSRLRIEDGKQKPKYPRRWILPPFL" complement(3660..3728) /exon /gene="T22E16.20" /number=1 /intron complement(3729..3820) /gene="T22E16.20" /number=1 /exon complement(3821..4575) /gene="T22E16.20" /number=2 /intron complement(4576..5037) /gene="T22E16.20" /number=2 /exon complement(5038..5077) /gene="T22E16.20" /number=3 /intron complement(5078..5261) /gene="T22E16.20" /number=3 /exon complement(5262..5330) /gene="T22E16.20" /number=4 /gene="T22E16.30" 10334..11350 /gene="T22E16.30" join(10334..11042,11136..11350) CDS /gene="T22E16.30"</pre>	<pre>/note="EMBL:AF155818 most likely is not full length, therefore N-terminal stretch" /codon_start=1 /product="zinc finger protein OBP3" /protein_id="CAB75895.1" /db_xref="GI:7076780" /db_xref="GOA:Q9M2U1" /db_xref="UniProt/Swiss-Prot:Q9M2U1" /translation="WQGNQHLECVTTDQNNYVLRQLSSPPTSQVAGSSQARVNSM VERARIAKVPPEALACPCDCSTNTKECYFNNYSLTOPRHFCKTKRYVWTRGGSLRN VPVGGFRNRKRSKRSKSTVVVSTNDTNTSSITSRSYSNPSPKFSYGOIPEPNSN LPILPQLSGDYNSNTGLDFGGTQISNMSGSSSGGILDAMKRIIPSQQAQPPFL INTTGLQSSNALYPLLEGGSATQTRNVKAEENDQDRGDDGVDGNNLSRNPFLGNINI NSGRNEEYTSWGGSSWTGFTSNNSTGHLSP" 10334..11042 /gene="T22E16.30" /number=1 /intron 11043..11135 /gene="T22E16.30" /number=1 /exon 11136..11350 /gene="T22E16.30" /number=2 /gene="T22E16.40" 13994..15369 join(13994..14039,14159..14261,14513..14610,14695..14795, 15214..15369) /gene="T22E16.40" /note="Contains Ubiquitin-conjugating enzymes active site AA79-93" /codon_start=1 /product="ubiquitin-conjugating enzyme UBC3" /protein_id="CAB75896.1" /db_xref="GI:7076781" /db_xref="GOA:P42747" /db_xref="InterPro:IPR000608" /db_xref="UniProt/Swiss-Prot:P42747" /translation="MANNQASLLQKLDLCKPVDGFSAGLVDEKRVFQWSVIMG PDDTLVEGFFNAIMSPFENYVPSPTVTFTSEMHVNPVYSDGKVCISILHHPGDDPH GYELASERTVPVHTVESIVLSIISMLSGPNDESPANVEAAKEWRDNRRAEFRKVSRCV RSQEML" 13994..14039 /gene="T22E16.40" /number=1 /intron 14040..14158 /gene="T22E16.40" /number=1 /exon 14159..14261 /gene="T22E16.40" /number=2 /intron 14262..14512 /gene="T22E16.40" /number=2 /exon 14513..14610 /gene="T22E16.40" /number=3 /intron 14611..14694 /gene="T22E16.40" /number=3 /exon 14695..14795 /gene="T22E16.40" /number=4 /intron 14796..15213 /gene="T22E16.40" /number=4 /exon 15214..15369 /gene="T22E16.40" /number=5 /gene="T22E16.50" 17259..18094 /gene="T22E16.50" join(17259..17479,17564..17699,17867..18094) /gene="T22E16.50" /note="similarity to MtN24, Medicago truncatula,"</pre>
--	---	--



EMBL:MTV15290  
Contains Prokaryotic membrane lipoprotein lipid attachment  
site AAI64-174  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB75897.1"  
/db\_xref="GI:7076782"  
/db\_xref="InterPro:IPR006459"  
/db\_xref="UniProt/TREMBL:Q9M2U0"  
/translation="NRSHPAENGSPRLRDTHFHSVTAOKLRRFNSLIILLRLAS  
FSFSLASAVFMTNRSHPAENGSPRLRDTHFHSVTAOKLRRFNSLIILLRLAS  
TULPAPQWPFQHDQVPSYLLSAGSNAALARTMRGGDCTANKAFCLQSDVAG  
LGFAAPFLFAPSSCSGFRVACFLITGSRFHDYS"  
17259..17479  
exon

Query Match	80.9%;	Score 6131.2;	DB 15;	Length 103240;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 6136;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;
QY	1	AGCTCTATTAAATCAAGAGAGCAGCAAAATAAGCAAAAACTCAAAACCTTAAGTTTCTGA	60	
DB	6232	AGCTCTATTAAATCAAGAGAGCAGCAAAATAAGCAAAAACTCAAAACCTTAAGTTTCTGA	6291	
QY	61	ATATGAAGGGGTAGATAATCATCTCTCAACTAGTTAAAAAAGTAATGATAAAATTA	120	
DB	6292	ATATGAAGGGGTAGATAATCATCTCTCAACTAGTTAAAAAAGTAATGATAAAATTA	6351	
QY	121	AAAAACAATGNTCAATTAAGAGACAGTAGTTTATGATATATATGTTGGGATCGATTA	180	
DB	6352	AAAAACAATGNTCAATTAAGAGACAGTAGTTTATGATATATATGTTGGGATCGATTA	6411	
QY	181	GTTGACATCAAAAGATCAAAATATGACGGTAAATGCCCCAAATACCAAGGCGCAAT	240	
DB	6412	GTTGACATCAAAAGATCAAAATATGACGGTAAATGCCCCAAATACCAAGGCGCAAT	6471	
QY	241	CATGGATATCAAAACCTTAATGTCATCAAAATCTATACTAACTATCTCAATTA	300	
DB	6472	CATGGATATCAAAACCTTAATGTCATCAAAATCTATACTAACTATCTCAATTA	6531	
QY	301	CAAGCTAGTAGATTTTCTCAAGTGAGCAACAGTCTAATCTCTGGAAGAAACCTGTT	360	
DB	6532	CAAGCTAGTAGATTTTCTCAAGTGAGCAACAGTCTAATCTCTGGAAGAAACCTGTT	6591	
QY	361	TTTTCCGAGTGAAATATCAATCTACTTACTACAAATTTGAGCATTAATCTAGTTTCT	420	
DB	6592	TTTTCCGAGTGAAATATCAATCTACTTACTACAAATTTGAGCATTAATCTAGTTTCT	6651	
QY	421	CTGCAACTTTAAACACAGGTGTCAAGTGTCAACACCAGATCTAGCGTAAACACCTAAG	480	
DB	6652	CTGCAACTTTAAACACAGGTGTCAAGTGTCAACACCAGATCTAGCGTAAACACCTAAG	6711	
QY	481	CTAGTACTTAATAGATTAATGCTATTTTTCGACCAATTAATTAATTAATTAATTAAT	540	
DB	6712	CTAGTACTTAATAGATTAATGCTATTTTTCGACCAATTAATTAATTAATTAATTAAT	6771	
QY	541	TTTCGTGAAAGAAATGCCATGTTCTTCTGGAATCTTAGGAATCTTCTTACTAAAGATTG	600	
DB	6772	TTTCGTGAAAGAAATGCCATGTTCTTCTGGAATCTTAGGAATCTTCTTACTAAAGATTG	6831	
QY	601	GCATGCACCTGGTAAACGATTTCCATTTGTATATCTATCTCGATATTTTCCACCTTAATAT	660	
DB	6832	GCATGCACCTGGTAAACGATTTCCATTTGTATATCTATCTCGATATTTTCCACCTTAATAT	6891	
QY	661	CTTGAAGAAATTAAGATCAAAATTAATAGAGAACATATTTGTATCAATTTGTTAAATAGCCT	720	
DB	6892	CTTGAAGAAATTAAGATCAAAATTAATAGAGAACATATTTGTATCAATTTGTTAAATAGCCT	6951	
QY	721	TATATAGTCTGTAGATGTCGATCTATCGATCAAAAAACAATAGGAAGAAAGAGCTGCAG	780	
DB	6952	TATATAGTCTGTAGATGTCGATCTATCGATCAAAAAACAATAGGAAGAAAGAGCTGCAG	7011	
QY	781	AGAAGGTGTTGCTTTTACAACTATGAATATGGGCTCATGAGATGTACACTACAGATAAGCG	840	

DB	7012	AGAAGGTGTTGCTTTTACAACTATGAATATGGGCTCATGAGATGTACACTACAGATAAGCG	7071	
QY	841	AAGATTCTGTTGTCATGAAAAATGTTACTAAATAAAAAAACAATGCACTTCTATATA	900	
DB	7072	AAGATTCTGTTGTCATGAAAAATGTTACTAAATAAAAAAACAATGCACTTCTATATA	7131	
QY	901	AGACGAATTTCTTTTAAATTAATAATTTCTATAACAATAAAGATAAGTCTCCTTTAAA	960	
DB	7132	AGACGAATTTCTTTTAAATTAATAATTTCTATAACAATAAAGATAAGTCTCCTTTAAA	7191	
QY	961	AACATGCAAAAGAAATATATAGATTTACCGTATCAGATTTCATACAATTTTATATATTT	1020	
DB	7192	AACATGCAAAAGAAATATATAGATTTACCGTATCAGATTTCATACAATTTTATATATTT	7251	
QY	1021	TCAGCTTGAAAGATTAACAATGATCAAACTGTATCGTGTGTCCTCTATTCACCCCTAGA	1080	
DB	7252	TCAGCTTGAAAGATTAACAATGATCAAACTGTATCGTGTGTCCTCTATTCACCCCTAGA	7311	
QY	1081	AGAAAGTGAACATGGAACCTTTATGTTATGTCATACGGCGAGCTAGCTTCTCCCTACTTG	1140	
DB	7312	AGAAAGTGAACATGGAACCTTTATGTTATGTCATACGGCGAGCTAGCTTCTCCCTACTTG	7371	
QY	1141	TCCAATAGATGAAGACATTTATCACTCAGCTTCAGCTACTTCGAAGGCGCAACATATCGACA	1200	
DB	7372	TCCAATAGATGAAGACATTTATCACTCAGCTTCAGCTACTTCGAAGGCGCAACATATCGACA	7431	
QY	1201	AAAAATCGTTTTPAGCTCTATCATCTGTCTTTTGAAGAAAAATATCAACATATCAAAATACA	1260	
DB	7432	AAAAATCGTTTTPAGCTCTATCATCTGTCTTTTGAAGAAAAATATCAACATATCAAAATACA	7491	
QY	1261	TATACACACTCCCAACAATAATATCAACAATATATATGTTTACTACGAAATTCGAAT	1320	
DB	7492	TATACACACTCCCAACAATAATATCAACAATATATATGTTTACTACGAAATTCGAAT	7551	
QY	1321	GATATGCTCTTTGAAACCAACTAACTGTGAATTAACAAGTGAAGCAATATATCTATAT	1380	
DB	7552	GATATGCTCTTTGAAACCAACTAACTGTGAATTAACAAGTGAAGCAATATATCTATAT	7611	
QY	1381	ATTCTTTTCTTGTGTCTCAACTTCATCCCTTTCTAAGTAAATTCAAATTAATTTGGAAGTTT	1440	
DB	7612	ATTCTTTTCTTGTGTCTCAACTTCATCCCTTTCTAAGTAAATTCAAATTAATTTGGAAGTTT	7671	
QY	1441	TGTCATCTAAATTTGAAGTTCTTTTACCGGATCATTTTGTCTCAGGTTCATATATATCTTT	1500	
DB	7672	TGTCATCTAAATTTGAAGTTCTTTTACCGGATCATTTTGTCTCAGGTTCATATATATCTTT	7731	
QY	1501	CTTAGTCTGATACGAAACCTTATAATATATGATTAGAGAGAGACATGTTTGAATGTTA	1560	
DB	7732	CTTAGTCTGATACGAAACCTTATAATATATGATTAGAGAGAGACATGTTTGAATGTTA	7791	
QY	1561	TATTTTCTGTTAAACCACTCTTGATTTATGAATATATATAGTTAGGATATGTTTGAATGT	1620	
DB	7792	TATTTTCTGTTAAACCACTCTTGATTTATGAATATATATAGTTAGGATATGTTTGAATGT	7851	
QY	1621	TGTTGTGACATAGTGAAGTCCATTTAAAGAAAGTCTGATTTAAATTTTACGTTTGGAC	1680	
DB	7852	TGTTGTGACATAGTGAAGTCCATTTAAAGAAAGTCTGATTTAAATTTTACGTTTGGAC	7911	
QY	1681	CACAAATCTTTCTTTTAGAAATCGGGACTGGGACCTTCTCAACAACATGTCGGTCTTTT	1740	
DB	7912	CACAAATCTTTCTTTTAGAAATCGGGACTGGGACCTTCTCAACAACATGTCGGTCTTTT	7971	
QY	1741	ACTTAATCTTACGTACCCCTCACAATTCGTAAACATAAAATCATCAAAATAATATAGAGAC	1800	
DB	7972	ACTTAATCTTACGTACCCCTCACAATTCGTAAACATAAAATCATCAAAATAATATAGAGAC	8031	
QY	1801	TGGTGATCATATTCGAAATATTTTTCATTAATCAATGTTATCGGTAAAGTTATATAGG	1860	
DB	8032	TGGTGATCATATTCGAAATATTTTTCATTAATCAATGTTATCGGTAAAGTTATATAGG	8091	
QY	1861	GGTATTAACATCAAGATTCAGAAAGATTTAAAAACACCTTGTCTCGAATCATGATTTGAT	1920	

Db	8092	GGTATAACATCAAGAAATCACGAAAGAAATTAATAAAACAAACCTTGTGCAATCATGATTTGATT	8151
Qy	1921	TTTTGGCTTATATCTTTCTAAATTTTATATCTTTGTCTGCAAAATTAGCAACCAAAATATATA	1980
Db	8152	TTTTGGCTTATATCTTTCTAAATTTTATATCTTTGTCTGCAAAATTAGCAACCAAAATATATA	8211
Qy	1981	TATCTCTCTTCAACATCGAATCTTTATTTTGTGTAAGGCATTTTTTCTTAACAGA	2040
Db	8212	TATCTCTCTTCAACATCGAATCTTTATTTTGTGTAAGGCATTTTTTCTTAACAGA	8271
Qy	2041	GGAAATTTTACATCACTTTAGACTGCAATCTTCGGGATATAAAATCTCGCCATGCAAGGT	2100
Db	8272	GGAAATTTTACATCACTTTAGACTGCAATCTTCGGGATATAAAATCTCGCCATGCAAGGT	8331
Qy	2101	AATTTATTTTTCATGACAAAAAGCCAAATGGCGATAAATTAATACTATACTATGCAA	2160
Db	8332	AATTTATTTTTCATGACAAAAAGCCAAATGGCGATAAATTAATACTATACTATGCAA	8391
Qy	2161	AACGAAACTTTACTTTGGTTCATACCGAGGAAACAAAGGTACACTCGATTGTGCAACTC	2220
Db	8392	AACGAAACTTTTACTTTGGTTCATACCGAGGAAACAAAGGTACACTCGATTGTGCAACTC	8451
Qy	2221	CACCAAGACCAACCCACTACCAATTCACCTTTATTTGTTCTTTTATTTCACTCAAAATCT	2280
Db	8452	CACCAAGACCAACCCACTACCAATTCACCTTTATTTGTTCTTTTATTTCACTCAAAATCT	8511
Qy	2281	TTTTTAATTTTAAATTAATTCGAATTTATTCGCTTCTCTCGTTGTTTTTAACTTTTAAT	2340
Db	8512	TTTTTAATTTTAAATTAATTCGAATTTATTCGCTTCTCTCGTTGTTTTTAACTTTTAAT	8571
Qy	2341	TAAAAATTTGAAAGGAGTCCCTAGGGTTCTCTCTCTGCAATGGCCACTCTCGCTCTTCAC	2400
Db	8572	TAAAAATTTGAAAGGAGTCCCTAGGGTTCTCTCTCTGCAATGGCCACTCTCGCTCTTCAC	8631
Qy	2401	ATCTTTTGTGGGCACTATGTAAGTATGCAAAAAAAGGTAAGGTAAGGTAAGGTAAGGTAAGG	2460
Db	8632	ATCTTTTGTGGGCACTATGTAAGTATGCAAAAAAAGGTAAGGTAAGGTAAGGTAAGGTAAGG	8691
Qy	2461	TACATGCTAAGAAAACTCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2520
Db	8692	TACATGCTAAGAAAACTCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	8751
Qy	2521	TTATTTTGTGCAAAATTTATCGTTGCTGATGAAATAACAATATAGGTAAGGTAAGGTAAGG	2580
Db	8752	TTATTTTGTGCAAAATTTATCGTTGCTGATGAAATAACAATATAGGTAAGGTAAGGTAAGG	8811
Qy	2581	ATCGGCAAGGAGTGAATTTATCTACCACTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2640
Db	8812	ATCGGCAAGGAGTGAATTTATCTACCACTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	8871
Qy	2641	AACATACACAATAAACAATATGCAATTAATTTCCATCTTCAACCAAAATGTTGTAAGTGA	2700
Db	8872	AACATACACAATAAACAATATGCAATTAATTTCCATCTTCAACCAAAATGTTGTAAGTGA	8931
Qy	2701	AGTAATTTGATCTATGTTATACGCAATCTACTAAATTTTGGAAAAAAGAAAAA	2760
Db	8932	AGTAATTTGATCTATGTTATACGCAATCTACTAAATTTTGGAAAAAAGAAAAA	8991
Qy	2761	AACTATCTATACATGCTCCAAATTTATCTGCTTATGTAATTTATGCTGATATATAGAG	2820
Db	8992	AACTATCTATACATGCTCCAAATTTATCTGCTTATGTAATTTATGCTGATATATAGAG	9051
Qy	2821	ATGTTGGTGTCTTCTGAAATTTGATATATGTTCTTTTATCTCTGAAATATGATATGTA	2880
Db	9052	ATGTTGGTGTCTTCTGAAATTTGATATATGTTCTTTTATCTCTGAAATATGATATGTA	9111
Qy	2881	ATCATCATAGCAATTTTCAGGTTTAAACCTTATCGAATTCGCAAAAGATTGAAACCAAT	2940
Db	9112	ATCATCATAGCAATTTTCAGGTTTAAACCTTATCGAATTCGCAAAAGATTGAAACCAAT	9171
Qy	2941	TGGAAAAAATAAATAAGTTTATCTATTTTCTAAGCAGCGGTGAAATTCAGTATAA	3000
Db	9172	TGGAAAAAATAAATAAGTTTATCTATTTTCTAAGCAGCGGTGAAATTCAGTATAA	9231

Qy	3001	TATATGAAACAAATTTCTTACCTGATATAATATGTATACAGTGATACAAACCAAGACGATCA	3060
Db	9232	TATATGAAACAAATTTCTTACCTGATATAATATGTATACAGTGATACAAACCAAGACGATCA	9291
Qy	3061	CGTATATGATTTGACTTGCACAAAATAAGCAAAACAAATACCTGTTTCAAAATCGACCTTAAT	3120
Db	9292	CGTATATGATTTGACTTGCACAAAATAAGCAAAACAAATACCTGTTTCAAAATCGACCTTAAT	9351
Qy	3121	TCCAAAAAGGTTAGTATAAGTAAGAGGCTTTTATTTATGAAACCAAAAGAAATAAG	3180
Db	9352	TCCAAAAAGGTTAGTATAAGTAAGAGGCTTTTATTTATGAAACCAAAAGAAATAAG	9411
Qy	3181	AGCCTAAGAGAAATGATGAAATTCGAAACAGAAAAAGAGCAATGTTATAGAAAAAGAAAA	3240
Db	9412	AGCCTAAGAGAAATGATGAAATTCGAAACAGAAAAAGAGCAATGTTATAGAAAAAGAAAA	9471
Qy	3241	AAAGAGAGATAAGAGAAATTAAGAAAAACAATAAATTAACAAAGGAACTTCAATTTCT	3300
Db	9472	AAAGAGAGATAAGAGAAATTAAGAAAAACAATAAATTAACAAAGGAACTTCAATTTCT	9531
Qy	3301	TCTCTTTATCCATTCAGCTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3360
Db	9532	TCTCTTTATCCATTCAGCTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	9591
Qy	3361	CAATTTCTTTCTTATGATGATATCCACCATATCTCGGACCTCTTACCTTAAAGGA	3420
Db	9592	CAATTTCTTTCTTATGATGATATCCACCATATCTCGGACCTCTTACCTTAAAGGA	9651
Qy	3421	TACAAATTAAGAGATTCAGGATGTTTCTCATCTCTCTCCAGTGAATCAGTTTCGATTCCT	3480
Db	9652	TACAAATTAAGAGATTCAGGATGTTTCTCATCTCTCTCCAGTGAATCAGTTTCGATTCCT	9711
Qy	3481	AAAATTTGCGAGGTAAGAAATTCAGTTATGATATTTGCTAGATGTTTCTGATTCGTTCC	3540
Db	9712	AAAATTTGCGAGGTAAGAAATTCAGTTATGATATTTGCTAGATGTTTCTGATTCGTTCC	9771
Qy	3541	TTTTTCTTCCAGCTCGATCAAGATTTTATGAAATTTGATGAGATTTTCTCGACAAAT	3600
Db	9772	TTTTTCTTCCAGCTCGATCAAGATTTTATGAAATTTGATGAGATTTTCTCGACAAAT	9831
Qy	3601	TCCTAGCTATTGTGCGCGCATATATATTACTTATGAATTTCTTAGTGTATTAACCC	3660
Db	9832	TCCTAGCTATTGTGCGCGCATATATATTACTTATGAATTTCTTAGTGTATTAACCC	9891
Qy	3661	TTTTTTTTTCTGCTCTCTCGAATATACGAAATATATGAAGATGATTTCAATTTTGGTC	3720
Db	9892	TTTTTTTTTCTGCTCTCTCGAATATATGAAGATGATTTCAATTTTGGTC	9951
Qy	3721	TTTTTTTTTCTGCTCTGAGCTTTTAAATAATTTCTTAGTTGATAAAAACTTTTCT	3780
Db	9952	TTTTTTTTTCTGCTCTGAGCTTTTAAATAATTTCTTAGTTGATAAAAACTTTTCT	10011
Qy	3781	TGCTCTTCTCAAGGCTTATGATATATGTTTTTCTTACAGGATTAATTTTCTCTTGGTT	3840
Db	10012	TGCTCTTCTCAAGGCTTATGATATATGTTTTTCTTACAGGATTAATTTTCTCTTGGTT	10071
Qy	3841	AGATTTTATACCGCGCATGGAATATFCACTTCAAAAAATAAAAGTTTAAAGTACTATG	3900
Db	10072	AGATTTTATACCGCGCATGGAATATFCACTTCAAAAAATAAAAGTTTAAAGTACTATG	10131
Qy	3901	ACTTTAACTGAGTTATTTATCCATTTCTTTTTCAGCTTTTTCAGCTTTTTCAGCTTTTTC	3960
Db	10132	ACTTTAACTGAGTTATTTATCCATTTCTTTTTCAGCTTTTTCAGCTTTTTCAGCTTTTTC	10191
Qy	3961	ATCTGCAATTTCTGTCAAAAGTAGTCACAATTTTATCTATTTTCTCTCCGACCAA	4020
Db	10192	ATCTGCAATTTCTGTCAAAAGTAGTCACAATTTTATCTATTTTCTCTCCGACCAA	10251
Qy	4021	TGTTTCAAACTCGAATTCCTTTTCAAAAGTTTCTGCTTTATATATAAAGTCTGAACT	4080
Db	10252	TGTTTCAAACTCGAATTCCTTTTCAAAAGTTTCTGCTTTATATATAAAGTCTGAACT	10311

QY 4081 AATTAGTACAAATTATGTTAATATGCGACGAGGAAACCAACATCAGCTAGAAATGTGTAC 4140  
DB 10312 AATTAGTACAAATTATGTTAATATGCGACGAGGAAACCAACATCAGCTAGAAATGTGTAC 10371  
QY 4141 AACTGACAGAACCCCTAAATAATCTTACGGGAGCTCTCATCAACACCGACTTCTCAGGT 4200  
DB 10372 AACTGACAGAACCCCTAAATAATCTTACGGGAGCTCTCATCAACACCGACTTCTCAGGT 10431  
QY 4201 TCGAGGTTCCGATCGATCAGCTAGAGTGAATTCATATGTTGGGAACGTGCTCGGATCGCAAAAGT 4260  
DB 10432 TCGAGGTTCCGATCGATCAGCTAGAGTGAATTCATATGTTGGGAACGTGCTCGGATCGCAAAAGT 10491  
QY 4261 CCCATTGCTGAGCAGCTCTAAATTCGCCCTAGATGTGACTCAACCAATACTAAGTTCTG 4320  
DB 10492 CCCATTGCTGAGCAGCTCTAAATTCGCCCTAGATGTGACTCAACCAATACTAAGTTCTG 10551  
QY 4321 TTACTTCAATAACTATAGCGCTTACTCAACCTCGCCATTTCTGCAAAACATGTCGTGCTA 4380  
DB 10552 TTACTTCAATAACTATAGCGCTTACTCAACCTCGCCATTTCTGCAAAACATGTCGTGCTA 10611  
QY 4381 TTGGACACGTGCGGTTCTTCCAGGAATGTTCTGTTGGAGGAGGCTTTAGGAGGAACAA 4440  
DB 10612 TTGGACACGTGCGGTTCTTCCAGGAATGTTCTGTTGGAGGAGGCTTTAGGAGGAACAA 10671  
QY 4441 GAGAAGCAATCCAGATCGAATCTACGCTGCTGCTCGACTGATATACTACTAGTAC 4500  
DB 10672 GAGAAGCAATCCAGATCGAATCTACGCTGCTGCTCGACTGATATACTACTAGTAC 10731  
QY 4501 TTCACTACCTTACTTCTCGCCCAAGTTACTCAAAACCTGACCAAGTTTCATAGCTACGGTCA 4560  
DB 10732 TTCACTACCTTACTTCTCGCCCAAGTTACTCAAAACCTGACCAAGTTTCATAGCTACGGTCA 10791  
QY 4561 AATCCCGGAGTTTAAATTCCAACTTGGCCACTTGTGCTCTCTCCAAAGCCTTGGAGATTA 4620  
DB 10792 AATCCCGGAGTTTAAATTCCAACTTGGCCACTTGTGCTCTCTCCAAAGCCTTGGAGATTA 10851  
QY 4621 CAATTCAGCAACACCTGGATAGATTTTGGTGGAACTCAAAATAGCAACATGATAGTGG 4680  
DB 10852 CAATTCAGCAACACCTGGATAGATTTTGGTGGAACTCAAAATAGCAACATGATAGTGG 10911  
QY 4681 TATGAGTCTTCTAGTGTGGGATCTTGGATGCATGGAGATACCTCCATCAACAAGCTCA 4740  
DB 10912 TATGAGTCTTCTAGTGTGGGATCTTGGATGCATGGAGATACCTCCATCAACAAGCTCA 10971  
QY 4741 GCAATTCCTCTTCTGATCAACACTACCGGATTTGGTGGCAATCTTCAAGCGGTTATATCC 4800  
DB 10972 GCAATTCCTCTTCTGATCAACACTACCGGATTTGGTGGCAATCTTCAAGCGGTTATATCC 11031  
QY 4801 ATTACTAGAGTAAAGGAGGTGTTAATCAAGGTGATTTCTCAAGAGAGTAGTGATTA 4860  
DB 11032 ATTACTAGAGTAAAGGAGGTGTTAATCAAGGTGATTTCTCAAGAGAGTAGTGATTA 11091  
QY 4861 TTCCCAATCAGCTAAATGTTTAAAGCCCTTGAATGATTTTCTCAGGCGGGTTAGGCCAC 4920  
DB 11092 TTCCCAATCAGCTAAATGTTTAAAGCCCTTGAATGATTTTCTCAGGCGGGTTAGGCCAC 11151  
QY 4921 GCAAAACAGAAATGTGAAGCGGGAAGAAATGATCAGGATCGGGGTAGGGATGGGATGG 4980  
DB 11152 GCAAAACAGAAATGTGAAGCGGGAAGAAATGATCAGGATCGGGGTAGGGATGGGATGG 11211  
QY 4981 AGTGAATAACTTATCAAGAAACTTTTGGGTAATATCAACATAAACTCAGGACGGAACGA 5040  
DB 11212 AGTGAATAACTTATCAAGAAACTTTTGGGTAATATCAACATAAACTCAGGACGGAACGA 11271  
QY 5041 GGAATACACATCATGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCCAACATCAAC 5100  
DB 11272 GGAATACACATCATGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCCAACATCAAC 11331  
QY 5101 AGGCCATCTCTCAATCTAAGTACTCAGCACTAGCTATTCTTGATGATTTCTTTGTTGGTT 5160  
DB 11332 AGGCCATCTCTCAATCTAAGTACTCAGCACTAGCTATTCTTGATGATTTCTTTGTTGGTT 11391  
QY 5161 GGGGTGTACATTTGGTGTGTCATCGGAGTTATTGCTGAGGAAGATCAAAACCATCGAGCT 5220

DB 11392 GGGGTGTACATTTGGTGTCTTGTCTGCGAGTTATTGCTGAGGAAGATCAAAACCATGAGCT 11451  
QY 5221 ATATCCAAAGGCTAAATTTTGGAGCTCAAGGAAAGGTATGGTTATAAAACTATCTTTTGG 5280  
DB 11452 ATATCCAAAGGCTAAATTTTGGAGCTCAAGGAAAGGTATGGTTATAAAACTATCTTTTGG 11511  
QY 5281 ATCTTTTAAAGATCTTCAAGGTGTAGTATGTTTATGGTTGGCTTCTGGTGATATTTA 5340  
DB 11512 ATCTTTTAAAGATCTTCAAGGTGTAGTATGTTTATGGTTGGCTTCTGGTGATATTTA 11571  
QY 5341 TGTGTTTATTAAGATTTGTTCTTATATATTTGGCTATATATAGAGGTGTGGTGATATGAT 5400  
DB 11572 TGTGTTTATTAAGATTTGTTCTTATATATTTGGCTATATATAGAGGTGTGGTGATATGAT 11631  
QY 5401 GAAATCAAGAGTTGATGTTTGGAAACTTTTTGTTGTTTCAATGAATAATCATCGAATTTCT 5460  
DB 11632 GAAATCAAGAGTTGATGTTTGGAAACTTTTTGTTGTTTCAATGAATAATCATCGAATTTCT 11691  
QY 5461 CAAATTTCTGGAGACCCATTTATGAGACATCTATAGAACATATATGTAATGTA 5520  
DB 11692 CAAATTTCTGGAGACCCATTTATGAGACATCTATAGAACATATATGTAATGTA 11751  
QY 5521 TATTAACCGTACTTAAGTCGAATTTTATGACCAAGTAAATAAAATATGCGCGAATGTACA 5580  
DB 11752 TATTAACCGTACTTAAGTCGAATTTTATGACCAAGTAAATAAAATATGCGCGAATGTACA 11811  
QY 5581 TGTCAATATCGAGTTTAAACTATTTTTCCCAATATAAACAATATTTCTCTTTTCGTCCTAA 5640  
DB 11812 TGTCAATATCGAGTTTAAACTATTTTTCCCAATATAAACAATATTTCTCTTTTCGTCCTAA 11871  
QY 5641 CTATATATCTCTTATCTGATTTCTTATTTCTCTTTTAAATCTCTTTTTCCTTTCCCTCCAA 5700  
DB 11872 CTATATATCTCTTATCTGATTTCTTATTTCTCTTTTAAATCTCTTTTTCCTTTCCCTCCAA 11931  
QY 5701 GACACAAAAAAGGAAATACAGAAACGAAAGGAGATTTTAAAGATTCATAACCCAC 5760  
DB 11932 GACACAAAAAAGGAAATACAGAAACGAAAGGAGATTTTAAAGATTCATAACCCAC 11991  
QY 5761 GAGAAATATGCACTTAAATTCAGACTAATCCCAAAATTTCAAGAAATTTATGTAATTTTG 5820  
DB 11992 GAGAAATATGCACTTAAATTCAGACTAATCCCAAAATTTCAAGAAATTTATGTAATTTTG 12051  
QY 5821 CGATTAATAATGTTGTTTCAATCAATTAATGGCCAACTAACTAATTTGAAAGAACATGGAA 5880  
DB 12052 CGATTAATAATGTTGTTTCAATCAATTAATGGCCAACTAACTAATTTGAAAGAACATGGAA 12111  
QY 5881 TGACTGAAACCATGCAATATCTCTCAAGTCTCAACCTATGAAGAATCATGTAACCAATAG 5940  
DB 12112 TGACTGAAACCATGCAATATCTCTCAAGTCTCAACCTATGAAGAATCATGTAACCAATAG 12171  
QY 5941 ACTATCATCATGATTAATTAATGCAATGATCTATAATGTAATCTTTGAAACATAGATATGTC 6000  
DB 12172 ACTATCATCATGATTAATTAATGCAATGATCTATAATGTAATCTTTGAAACATAGATATGTC 12231  
QY 6001 ATTTATCTGGATATAAGATGCGGTTTAAACCTACTTTGCAATTTTGTATATCTTTCT 6060  
DB 12232 ATTTATCTGGATATAAGATGCGGTTTAAACCTACTTTGCAATTTTGTATATCTTTCT 12291  
QY 6061 TCTAATACATATGATCAATACACTTTTCTTTTAAAGAAATTTAAAGAACTTATTTTCAAC 6120  
DB 12292 TCTAATACATATGATCAATACACTTTTCTTTTAAAGAAATTTAAAGAACTTATTTTCAAC 12351  
QY 6121 ATCGATCAATTTTCTTTTGT 6144  
DB 12352 ATCGATCAACATGATGATGGT 12375

RESULT 2  
AF187951/c  
LOCUS  
DEFINITION Activation-tagging vector pSKI015, complete sequence.  
ACCESSION AF187951

10138 bp DNA linear SYN 11-MAY-2000

VERSION	AF187951.1	GI:6537289	
KEYWORDS	Activation-tagging vector pSKI015		
SOURCE	Activation-tagging vector pSKI015		
ORGANISM	other sequences; artificial sequences; vectors.		
REFERENCE	1 (bases 1 to 10138)		
AUTHORS	Weigel, D., Ahn, J. H., Blazquez, M. A., Borevitz, J. O., Kardailsky, I., Christensen, S. K., Fankhauser, C., Ferrandiz, C., Sack, S., Wang, Z., Malancharuvil, E. J., Neff, M. M., Nguyen, J. T., Sack, S., Wang, Z., Xia, Y., Dixon, R. A., Harrison, M. J., Lamb, C. J., Yanofsky, M. F. and Chory, J.		
TITLE	Activation tagging in Arabidopsis		
JOURNAL	Plant Physiol. 122 (4), 1003-1013 (2000)		
PUBMED	10759496		
REFERENCE	2 (bases 1 to 10138)		
AUTHORS	Kardailsky, I. and Weigel, D.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-SEP-1999) Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
FEATURES	Location/Qualifiers		
source	1..10138		
	/organism="Activation-tagging vector pSKI015"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:109189"		
misc_feature	12..31		
	/note="T7 RNA polymerase promoter"		
misc_feature	36..2204		
	/note="derived from pUC19"		
misc_feature	2205..2662		
	/note="derived from phage f1"		
misc_feature	2663..2829		
	/note="derived from pUC19"		
misc_feature	2830..2849		
	/note="T7 RNA polymerase promoter"		
misc_feature	2894..4178		
	/note="cauliflower mosaic virus 35S gene promoter enhancer sequences (tetramer)"		
misc_feature	4270..4294		
	/note="T-DNA right border"		
misc_feature	7690..7713		
	/note="T-DNA left border"		
mRNA	complement(8054..8784)		
	/product="octopine synthase"		
	/note="3' end"		
mRNA	complement(<8785..>9420)		
CDS	complement(8791..9342)		
	/note="promotes resistance to glutamine synthetase inhibitors"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="AAF15587.1"		
	/db_xref="GI:6537290"		
	/translation="MSPRRPADIRREADMPACTIVNHVIETSTVNFRTPEOPBQ EWTDLLRLRERYPHLVAEVDGEVAGIAYAGPKARNAYDWTAEISTVYVSPHQTL GSTLYTHLKSLEAGFSVAVIGLPNDPSVRMEALGAPRGLDRAAGFKHGNWHD VGFQWQDLSLPVPRPVLPVTEM"		
mRNA	complement(9421..10119)		
	/product="mannopine synthase"		
	/note="5' end and promoter"		
ORIGIN	Query Match 18.5%; Score 1402.4; DB 11; Length 10138;		
	Best Local Similarity 99.9%; Pred. No. 1.7e-195;		
	Matches 1403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	6177	TTGTTTCGATCTAGATATACATCAATCCATCTGCTTTTGAAGACGCTGTTGAACGCTCT	6236
DB	4261	TAGTTTCGATCTAGATATACATCAATCCATCTGCTTTTGAAGACGCTGTTGAACGCTCT	4202
QY	6237	TCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGGGTAGAG	6296
DB	4201	TCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGGGTAGAG	4142

QY 7377 TCTACTGTCCTTTCGATGAAGTGCACAGATAGCTGGCAATGCATCCGAGGAGTTTCCC 7436  
Db 3061 TCTACTGTCCTTTCGATGAAGTGCACAGATAGCTGGCAATGCATCCGAGGAGTTTCCC 3002  
QY 7437 GATATTACCTTTGTTGAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTG 7496  
Db 3001 GATATTACCTTTGTTGAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTG 2942  
QY 7497 ATATTCTTGGAGTAGACGAGAGTGTGCTCTCACCATGTTGGGAGTCCACTAGTTCTAG 7556  
Db 2941 ATATTCTTGGAGTAGACGAGAGTGTGCTCTCACCATGTTGGGAGTCCACTAGTTCTAG 2882  
QY 7557 AGCGGCGCCACCGGCTGGAGCT 7580  
Db 2881 AGCGGCGCCACCGGCTGGAGCT 2858

## RESULT 3

AF218466/c 10450 bp DNA circular SYN 11-MAY-2000  
LOCUS AF218466 Activation tagging vector pSKI074, complete sequence.  
DEFINITION AF218466  
ACCESSION AF218466  
VERSION AF218466.1 GI:6715465  
KEYWORDS  
SOURCE Activation tagging vector pSKI074  
ORGANISM Activation tagging vector pSKI074  
other sequences; artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 10450)  
AUTHORS Weigel, D., Ahn, J.H., Blazquez, M.A., Borevitz, J.O.,  
Christensen, S.K., Fankhauser, C., Ferrandiz, C., Kardailsky, I.,  
Malancharuvil, E.J., Neff, M.M., Nguyen, J.T., Sato, S., Wang, Z.,  
Xia, Y., Dixon, R.A., Harrison, M.J., Lamb, C.J., Yanofsky, M.F. and  
Chory, J.

TITLE Activation tagging in Arabidopsis  
JOURNAL Plant Physiol. 122 (4), 1003-1013 (2000)  
PUBMED 10759496  
REFERENCE 2 (bases 1 to 10450)  
AUTHORS Weigel, D., Blazquez, M.A., Borevitz, J., Christensen, S.K.,  
Fankhauser, C., Ferrandiz, C., Malancharuvil, E.J., Neff, M.M.,  
Nguyen, J.T., Sato, S., Xia, Y., Wang, Z., Dixon, R.A., Harrison, M.J.,  
Lamb, C.J., Yanofsky, M.F. and Chory, J.  
Direct Submission  
Submitted (22-DEC-1999) Plant Biology Laboratory, The Salk  
Institute for Biological Studies, 10010 N. Torrey Pines Road, La  
Jolla, CA 92037, USA

## FEATURES

source  
1..10450  
/organism="activation tagging vector pSKI074"  
/mol\_type="other DNA"  
/db\_xref="taxon:112458"  
12..31  
/note="from T7 RNA polymerase"  
36..2204  
/note="derived from pUC19"  
1662..2205  
/note="derived from phage f1"  
2663..2829  
/note="derived from pUC19"  
2830..2849  
/note="from T7 RNA polymerase"  
2894..4178  
/note="from cauliflower mosaic virus 35S gene promoter"  
4270..4294  
/note="T-DNA right border"  
8054..8784  
/note="T-DNA left border"  
8064..8749  
/note="includes mannopine synthase promoter and 5' UTR  
sequence"  
8756..9739  
/product="neomycin phosphotransferase"  
8771..9565  
/function="confers kanamycin resistance"

/codon\_start=1  
/product="neomycin phosphotransferase"  
/protein\_id="AAF26422.1"  
/db\_xref="GI:6715466"  
/translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGSDAAVFRLSAQR  
PVLFKTDLGSLNLEQDEAARLRLSLWATTTGVPAAVLDVVTEAGRDMLLGEVPGDIL  
LSSHLAPAEKVISIMADAMRRLTLDPATCPDQAKHRIERATRMEAGLVDQDDIDE  
EHQGLAPAEELFARLKARMPDGEDLVVTHGDACLPLNIMVENGFSGFIDGRLGVADRY  
QDIALATDIDAEELGGEWADRFVLVYGIAPDPSQRIAFIRLLDEFF"  
9746..10430  
/note="from mannopine synthase"

## terminator

## ORIGIN

Query Match 18.5%; Score 1402.4; DB 11; Length 10450;  
Best Local Similarity 99.9%; Pred. No. 1.7e-195;  
Matches 1403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 6177 TTGTTTCGGATCTAGATATCACATCCACTGCTTTGGAAGAGCTGGTTCGAAAGCTCT 6236  
Db 4261 TAGTTTCGGATCTAGATATCACATCCACTGCTTTGGAAGAGCTGGTTCGAAAGCTCT 4202  
QY 6237 TCCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAG 6296  
Db 4201 TCCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAG 4142  
QY 6297 GCATCTTGAACGATAGCCTTCTTTATCGCAATGATGCGCATTTGTAGAGCCATCTTCC 6356  
Db 4141 GCATCTTGAACGATAGCCTTCTTTATCGCAATGATGCGCATTTGTAGAGCCATCTTCC 4082  
QY 6357 TTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTT 6416  
Db 4081 TTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTT 4022  
QY 6417 CCCGATATTTACCTTTGTTGAAAGCTCTCAATPAGCCCTCTGGTCTTTCTGAGACTGTATCT 6476  
Db 4021 CCCGATATTTACCTTTGTTGAAAGCTCTCAATPAGCCCTCTGGTCTTTCTGAGACTGTATCT 3962  
QY 6477 TTGATATTTCTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGGATCTAGATATCA 6536  
Db 3961 TTGATATTTCTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGGATCTAGATATCA 3902  
QY 6537 CATCAATCCACTGCTTTTGAAGACGTGGTGAAGCTCTTCTTTTCCACGATGTTCCCTC 6596  
Db 3901 CATCAATCCACTGCTTTTGAAGACGTGGTGAAGCTCTTCTTTTCCACGATGTTCCCTC 3842  
QY 6597 GTGGGTGGGGTTCATCTTTGGGACCACTGTGCGTAGAGGCATCTTTGAACGATAGCCTTT 6656  
Db 3841 GTGGGTGGGGTTCATCTTTGGGACCACTGTGCGTAGAGGCATCTTTGAACGATAGCCTTT 3782  
QY 6657 CTTTATCGCAATGATGCGATTTGTAGAGCCATCTTCTCTTTTCTACTGCTCTTTCGATG 6716  
Db 3781 CTTTATCGCAATGATGCGATTTGTAGAGCCATCTTCTCTTTTCTACTGCTCTTTCGATG 3722  
QY 6717 AAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATTTACCTTTGTTGA 6776  
Db 3721 AAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATTTACCTTTGTTGA 3662  
QY 6777 AAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGAGC 6836  
Db 3661 AAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGAGC 3602  
QY 6837 AAGTGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTTGTCTTGA 6896  
Db 3601 AAGTGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTTGTCTTGA 3542  
QY 6897 GACGTGTTGGACGCTCTTCTTTTCCACGATGTTCTCGTGGGTGGGGTCCACTCTTGG 6956  
Db 3541 GACGTGTTGGACGCTCTTCTTTTCCACGATGTTCTCGTGGGTGGGGTCCACTCTTGG 3482  
QY 6957 GGACCACTGTGCGTAGAGGCATCTTTGAACGATAGCCTTTTCCCTTTATCGCAATGATGGCAT 7016  
Db 3481 GGACCACTGTGCGTAGAGGCATCTTTGAACGATAGCCTTTTCCCTTTATCGCAATGATGGCAT 3422



QY	7017	TTGTAGAGCCATCTTCCCTTTCTACTGTCCTTTTCATGAAAGTGACAGATAGCTGGCAA	7076
Db	3421	TTGTAGAGCCATCTTCCCTTTCTACTGTCCTTTTCATGAAAGTGACAGATAGCTGGCAA	3362
QY	7077	TGGAATCCGAGGAGGTTTCCCGATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGG	7136
Db	3361	TGGAATCCGAGGAGGTTTCCCGATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGG	3302
QY	7137	TCTTCTGAGACTGTATCTTTGATATCTTGGAGTAGACAGAGTGTCTGCTCCACCATG	7196
Db	3301	TCTTCTGAGACTGTATCTTTGATATCTTGGAGTAGACAGAGTGTCTGCTCCACCATG	3242
QY	7197	TTGGGGATCTAGATATCAATCAATCCACTTCTTTGAAGACGTGGTTGGAAAGCTTCTCT	7256
Db	3241	TTGGGGATCTAGATATCAATCAATCCACTTCTTTGAAGACGTGGTTGGAAAGCTTCTCT	3182
QY	7257	TTTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGCA	7316
Db	3181	TTTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGCA	3122
QY	7317	TCTTGAAAGATGCTTCTCTTTATCGCAATGATGCGCATTTGTAGAAGCCATCTTCCCTT	7376
Db	3121	TCTTGAAAGATGCTTCTCTTTATCGCAATGATGCGCATTTGTAGAAGCCATCTTCCCTT	3062
QY	7377	TCTACTGCTCTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCC	7436
Db	3061	TCTACTGCTCTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCC	3002
QY	7437	GATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTG	7496
Db	3001	GATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTG	2942
QY	7497	ATATTCTGGAGTAGACGAGAGTGTCTGCTCCACCATTTGGGATCCACTAGTCTTAG	7556
Db	2941	ATATTCTGGAGTAGACGAGAGTGTCTGCTCCACCATTTGGGATCCACTAGTCTTAG	2882
QY	7557	ACGGCCGCGCACCGCGGTGGAGCT	7580
Db	2881	ACGGCCGCGCACCGCGGTGGAGCT	2858
RESULT 4			
AK221402			
LOCUS			
DEFINITION	Arabidopsis thaliana gene for zinc finger protein OBP3, complete cds, clone: RAFL25-48-C17.		
ACCESSION	AK221402		
VERSION	AK221402.1		
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Totoki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K., Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K., Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
AUTHORS	Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs Unpublished		
JOURNAL	2 (bases 1 to 1457)		
REFERENCE	Totoki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K., Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K., Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (22-MAR-2005) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: msei@gs.riken.jp).		
REFERENCE	URL: http://range.gsc.riken.jp/, Tel: 81-45-503-9625,		

COMMENT	Fax: 81-45-503-9586) An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720; Seki et al. (2002) Science 296:141-145). This clone is in a modified pBluescript vector. Please visit our web site (http://range.gsc.riken.jp/) for further details.		
FEATURES	Location/Qualifiers		
source	1..1457		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="3"		
	/clone="RAFL25-48-C17"		
	/ecotype="Columbia"		
	/note="common name: thale cress"		
gene	1..1457		
	/gene="At3g55370"		
CDS	134..1105		
	/gene="At3g55370"		
	/codon_start=1		
	/product="zinc finger protein OBP3"		
	/protein_id="BAD94355.1"		
	/db_xref="GI:62320152"		
	/translation="MVFSLEPVQFDSQWQQNQHQLECVTTDQNPNNYLRLQSSP PTSOVAGSQAQVNSMVERARIKVPLEAALNCPKCDSTNTKFCFNNYSLTQPRHF CKTRVYWRGSLRNVPVGGGRNKRKSKRSKTSVVVTDNTSTSLTSRPSYSN PSKSHSYQIPEFNSNLPILPLOSIGDYSNLTGLDFGGTQISNMISSMSSGGILD AWRPPSOAQQFPFLINTTGLVQSSNALYELLEGCVSATQTRNKAEENDQDRDRD DQVNNLSRNLNINSGRNEEYTSWGNSSWTGFTSNNSGHLUSF"		
ORIGIN			
Query Match	15.3%; Score 1163; DB 15; Length 1457;		
Best Local Similarity	93.2%; Pred. No. 2.9e-160;		
Matches 1266; Conservative	0; Mismatches 0; Indels 93; Gaps 1;		
QY	4105	GCAGCAAGGGAACCAACATCAGCTAGAATGTGTCAACAACGACCAAGACCTTAATAATTA	4164
Db	184	GCAGCAAGGGAACCAACATCAGCTAGAATGTGTCAACAACGACCAAGACCTTAATAATTA	243
QY	4165	CTTACGGCAGCTCTCATCAACCACTTCTCAGGTTGAGGTTGAGTCAAGCTAGAGT	4224
Db	244	CTTACGGCAGCTCTCATCAACCACTTCTCAGGTTGAGGTTGAGTCAAGCTAGAGT	303
QY	4225	GAAATCAATGGTGGAAAGCTGCTCGATCGCAAAAGTCCCATTCCTGAAGCAGCTCTAAA	4284
Db	304	GAAATCAATGGTGGAAAGCTGCTCGATCGCAAAAGTCCCATTCCTGAAGCAGCTCTAAA	363
QY	4285	TTGCCCTAGATGTGACTCAACCAATACTAAAGTTCTGTACTTCAATAACTATAGCCTTAC	4344
Db	364	TTGCCCTAGATGTGACTCAACCAATACTAAAGTTCTGTACTTCAATAACTATAGCCTTAC	423
QY	4345	TCAACCTCGCCATTTCTGCAAAACATGTGCTCGCTATTGGACACGTTGGGGTTCCTTGAG	4404
Db	424	TCAACCTCGCCATTTCTGCAAAACATGTGCTCGCTATTGGACACGTTGGGGTTCCTTGAG	483
QY	4405	GAAATGTTCTGTGGAGGAGGCTTTAGGAGGAACAAGACAACCAATCCAGATCGAAATC	4464
Db	484	GAAATGTTCTGTGGAGGAGGCTTTAGGAGGAACAAGACAACCAATCCAGATCGAAATC	543
QY	4465	TACGGTCTGGTCTCGACTGATAACTACTAGTACTTCACTTCACTTCTCTCGCCCAAG	4524
Db	544	TACGGTCTGGTCTCGACTGATAACTACTAGTACTTCACTTCACTTCTCTCGCCCAAG	603
QY	4525	TTACTCAAAACCTTAGCAAGTTTCTAGCTACGGTCAAAATCCCGGAGTTTAAATCCAACTT	4584
Db	604	TTACTCAAAACCTTAGCAAGTTTCTAGCTACGGTCAAAATCCCGGAGTTTAAATCCAACTT	663
QY	4585	GCCCATCTTGGCTCTCTCTCAAGAGCTTGGAGATTACAAATCAAGCAACACTGGATTAGA	4644
Db	664	GCCCATCTTGGCTCTCTCTCAAGAGCTTGGAGATTACAAATCAAGCAACACTGGATTAGA	723
QY	4645	TTTTTGTGGAACTCAAAATAGCAACATGATAAGTGGTATAGTTCCTAGTGGTGGATCTTT	4704

```
Db 724 TTTTGTGGAATCAAAATAGCAACATGATAGTGTATGATCTTAGTGGTGGATCTT 783
Qy 4705 GGATGATGGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTCTTCTGATCAACAC 4764
Db 784 GGATGATGGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTCTTCTGATCAACAC 843
Qy 4765 TACCGGATGGTGCATCTTCAACCGCTTATATCAATCTAGTAAAGGTAAGGAGGTGT 4824
Db 844 TACCGGATGGTGCATCTTCAACCGCTTATATCAATCTAGTAAAGGTAAGGAGGTGT 888
Qy 4825 TAATCAAGGTGATCTTCAACGAGAGTAGTATTAATCAATCAGCTAATGTTTAAAGCC 4884
Db 889 ----- 888
Qy 4885 CTTGTGATTTTCTTTCAGGGGGGTAGCGCCAGCAACAAAGAAATGTGAAGCGGA 4944
Db 889 -----AGCGGGGTAGCGCCAGCAACAAAGAAATGTGAAGCGGA 930
Qy 4945 AGAAGATGATCAGGATCGGGGTAGGGATGGGATGGGATGGAATTAATCAAGAAATCTT 5004
Db 931 AGAAGATGATCAGGATCGGGGTAGGGATGGGATGGGATGGAATTAATCAAGAAATCTT 990
Qy 5005 TTTGGGTAAATCAACATAAACTCAGCGCAGGACGAGGAATACATCATCGGGAGGTAA 5064
Db 991 TTTGGGTAAATCAACATAAACTCAGCGCAGGACGAGGAATACATCATCGGGAGGTAA 1050
Qy 5065 CAGTCTTTCGACCGGTTTCACTCTCAACCAACTCAACAGGCCATCTCTCAATCTTAAGTACT 5124
Db 1051 CAGTCTTTCGACCGGTTTCACTCTCAACCAACTCAACAGGCCATCTCTCAATCTTAAGTACT 1110
Qy 5125 CAGCACTAGCTATCTTGTATGATCTTTTGTGGTGGGGTGTATCATTTGCTGTCAT 5184
Db 1111 CAGCACTAGCTATCTTGTATGATCTTTTGTGGTGGGGTGTATCATTTGCTGTCAT 1170
Qy 5185 GCGAGTATTGCTGAGGAGATCAACCACTGAGCTATATCAAGGCTAAATTTTGAGGC 5244
Db 1171 GCGAGTATTGCTGAGGAGATCAACCACTGAGCTATATCAAGGCTAAATTTTGAGGC 1230
Qy 5245 TCAAGGAAAGGTATGGTTATATAAACTATCTTTTGTATCTTTTAAAGATCTTCAAGTG 5304
Db 1231 TCAAGGAAAGGTATGGTTATATAAACTATCTTTTGTATCTTTTAAAGATCTTCAAGTG 1290
Qy 5305 TGAGTATGTTTATGGTGGCTTCTGGTGATATTAATGTTTATTAAGATTTGGCTTAT 5364
Db 1291 TGAGTATGTTTATGGTGGCTTCTGGTGATATTAATGTTTATTAAGATTTGGCTTAT 1350
Qy 5365 ATATTGGCTATATAGAGGTGGTGATATGATGATGATGATGATGATGATGATGATGATG 5424
Db 1351 ATATTGGCTATATAGAGGTGGTGATATGATGATGATGATGATGATGATGATGATGATG 1410
Qy 5425 CTTTCTTGTGCTTCAATGAATAATCATCGAATTTCTCA 5463
Db 1411 CTTTCTTGTGCTTCAATGAATAATCATCGAATTTCTCA 1449

RESULT 5
AF155818
LOCUS AF155818 1274 bp mRNA linear PLN 23-DEC-1999
DEFINITION Arabidopsis thaliana zinc finger protein OBP3 mRNA, complete cds.
ACCESSION AF155818
VERSION AF155818.1 GI:5059397
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1274)
REFERENCE
Kang,H.G. and Singh,K.B.
Characterization of Arabidopsis Dof Transcription Factors, a Novel
Zinc Finger Protein Family in Plants
JOURNAL Unpublished
```

```
REFERENCE 2 (bases 1 to 1274)
AUTHORS Kang,H.G. and Singh,K.B.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) MCDB, University of California, Los
Angeles, 405 Hilgard Ave., Los Angeles, CA 90095, USA
FEATURES
Location/Qualifiers
1..1274
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/ecotype="Columbia"
138..932
/note="transcriptional activator"
/codon_start=1
/product="zinc finger protein OBP3"
/protein_id="AAD38988.1"
/db_xref="GI:5059398"
/translation="MVERARIATAKPLPEALNCPDRDSTNTKFCYFNYSLTQPRHFC
KTCRYWTRGSLRNVPVGGFRNRKRSKSTVVTNDTNTTSTSLTSRPSYSNP
SKFHSYGGIPEFNSNLPILPLOSIGDYNSSNTGLDFTGTQISNMSGSSGGILDA
WRIPSOQAQQFPFLINTTGLVQSSNALYPLLEGVSATQTRNVKAEENDQKRGDGD
GVNLSRNLGNINISGRNBEYTSWGGNSSWTGFTSNSTGHLSP"
ORIGIN
Query Match 15.3%; Score 1156.2; DB 15; Length 1274;
Best Local Similarity 92.9%; Pred. No. 2.9e-159;
Matches 1261; Conservative 0; Mismatches 3; Indels 93; Gaps 1;
Qy 4105 GCAGCAAGGGAACCAACATCAGCTAGATGTGTCAACAAGTCCAGCAAGCCTTAATAATTA 4164
Db 11 GGAGTAAAGGAACCAACATCAGCTAGATGTGTCAACAAGTCCAGCAAGCCTTAATAATTA 70
Qy 4165 CTTACGGCAGCTCTCATCACCACCTTCTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGT 4224
Db 71 CTTACGGCAGCTCTCATCACCACCTTCTCAGGTTCAGGTTCAGGTTCAGGTTCAGGTTCAGT 130
Qy 4225 GAATTCATGTGGAAAGTGTCTCGATCGCAAAAGTCCCATTCCTGAGCAGCTCTAA 4284
Db 131 GAATTCATGTGGAAAGTGTCTCGATCGCAAAAGTCCCATTCCTGAGCAGCTCTAA 190
Qy 4285 TTGCCCTAGATGTGACTCAACCAATCTAAGTTCTGTACTTTCAATACTATAGCCTTAC 4344
Db 191 TTGCCCTAGATGTGACTCAACCAATCTAAGTTCTGTACTTTCAATACTATAGCCTTAC 250
Qy 4345 TCACCTCGCCATTTCTGCAAAACATGTCTGCTATTTGGACACGTCGGGTTCCTTGAG 4404
Db 251 TCACCTCGCCATTTCTGCAAAACATGTCTGCTATTTGGACACGTCGGGTTCCTTGAG 310
Qy 4405 GAATGTTCTGTGTGAGGAGGCTTTAGGAGGAACAAGAGAACAAATCCAGATCGAATC 4464
Db 311 GAATGTTCTGTGTGAGGAGGCTTTAGGAGGAACAAGAGAACAAATCCAGATCGAATC 370
Qy 4465 TACGGTCTGGTCTCGACTGATTAATCTAGTACTTCACTACCTTCTTCGCCCAAG 4524
Db 371 TACGGTCTGGTCTCGACTGATTAATCTAGTACTTCACTACCTTCTTCGCCCAAG 430
Qy 4525 TTACTCAAAACCTTAGCAAGTTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATCCAACCTT 4584
Db 431 TTACTCAAAACCTTAGCAAGTTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATCCAACCTT 490
Qy 4585 GCCCATCTTGCTCTCTCCAAAGCCTTGGAGATTCAATTAAGCAACACACCTGGATTAGA 4644
Db 491 GCCCATCTTGCTCTCTCCAAAGCCTTGGAGATTCAATTAAGCAACACACCTGGATTAGA 550
Qy 4645 TTTTGGTGGAACTCAATAAGCAACATGATAGTGTAGTCTAGTGGTGGATCTT 4704
Db 551 TTTTGGTGGAACTCAATAAGCAACATGATAGTGTAGTCTAGTGGTGGATCTT 610
Qy 4705 GGATGCATGGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTCTTCTTCTGATCAACAC 4764
Db 611 GGATGCATGGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTCTTCTTCTGATCAACAC 670
Qy 4765 TACCGGATGGTGCATCTTCAACCGCTTATATCAATCTAGAAGGTAAGGAGGTGT 4824
```



[illegible]

## RESULT 6

AF406991/c	AF406991	9663 bp	DNA	linear	SYN 13-NOV-2003
LOCUS					
DEFINITION	Tobacco rattle virus RNA2-based VIGS vector pTRV2, complete sequence.				

ACCESSION	AF406991		
VERSION	AF406991.1	GI:21389151	
KEYWORDS			
SOURCE	Tobacco rattle virus RNA2-based	VIGS vector	pYLL56
ORGANISM	Tobacco rattle virus RNA2-based	VIGS vector	pTRV2
	Other sequences:	artificial	sequences: vectors.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 9663)  
Liu, Y., Schiff, M., Marathe, R. and Dineash-Kumar, S.P.  
Tobacco Karl, EN51 and NPRL/NIM1 like genes are required for  
N-mediated resistance to tobacco mosaic virus  
Plant J. 30 (4), 415-429 (2002)  
12028572  
2 (bases 1 to 9663)  
Liu, Y.  
Direct Submission  
Submitted (08-AUG-2001) MCDB, Yale University, 165 Prospect Street,  
New Haven, CT 06520, USA

**FEATURES**  
**SOURCE**

## RESULT 7

```
AY788908/c
LOCUS       AY788908             12537 bp    DNA             circular SYN 02-MAR-2005
DEFINITION   N-terminal TAPA T-DNA vector pN-TAPA, complete sequence.
ACCESSION   AY788908
VERSION     AY788908.1   GI:55824365
KEYWORDS    'N-terminal TAPA T-DNA vector pN-TAPA'
SOURCE      N-terminal TAPA T-DNA vector pN-TAPA
ORGANISM    N-terminal TAPA T-DNA vector pN-TAPA
REFERENCE   1 (bases 1 to 12537)
AUTHORS     Rubio,V., Shen,Y., Saijo,Y., Liu,Y., Gusmaroli,G.,
            Dinesh-Kumar,S.P. and Deng,X.W.
TITLE       An alternative tandem affinity purification strategy applied to
            Arabidopsis protein complex isolation
JOURNAL     Plant J. 41 (5), 767-778 (2005)
PUBMED     15703063
REFERENCE   2 (bases 1 to 12537)
AUTHORS     Rubio,V. and Deng,X.W.
DIRECT SUBMISSION
TITLE       Submitted (21-OCT-2004) MCDB, Yale University, 165, Prospect St.,
            New Haven, CT 06511, USA
JOURNAL
FEATURES             Location/Qualifiers
     1..12537
     /organism="N-terminal TAPA T-DNA vector pN-TAPA"
     /mol_type="other DNA"
     /db_xref="taxon:300483"
     /note="derived from pPZP222"
     promoter
     1..769
     /note="duplicated CamV 35S promoter"
     misc_feature
     780..844
     /note="TMV UI omega sequence"
     misc_feature
     854..1231
     /note="2x protein A IgG binding domain"
     misc_feature
     1256..1279
     /note="protease 3C cleavage site"
     misc_feature
     1286..1303
     /note="6x HIS tag"
     misc_feature
     1325..1675
     /note="9x myc tag"
     misc_feature
     1704..3577
     /note="attR1-CmrR-cdb-attR2 GATEWAY cassette"
     terminator
     3607..3863
     /note="NOS terminator"
ORIGIN
Query Match      8.4%; Score 640.2; DB 11; Length 12537;
Best Local Similarity 98.4%; Pred. No. 3.6e-84;
Matches 660; Conservative 0; Mismatches 3; Indels 8; Gaps 1;

QY 6190 AGATATCATCAATCAATCCACTTCTTGAAGACGTGGTTGGAAACGTCTTCTTTTCCACGA 6249
DB 683 AGATATCATCAATCAATCCACTTCTTGAAGACGTGGTTGGAAACGTCTTCTTTTCCACGA 624
QY 6250 TGTTCCTCGTGGTGGGGTCCATCTTTGGGACCACTGCTGGTAGAGGCATCTTCCATGAA 6309
DB 623 TGTTCCTCGTGGTGGGGTCCATCTTTGGGACCACTGCTGGTAGAGGCATCTTCCATGAA 564
QY 6310 TAGCCCTTCTTTATCGCAATGATGGCAATTTGAGAACCATCTTCTTTTCTACTGTCC 6369
DB 563 TAGCCCTTCTTTATCGCAATGATGGCAATTTGAGAACCATCTTCTTTTCTACTGTCC 504
QY 6370 TTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATATCCC 6429
DB 503 TTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATATCCC 444
QY 6430 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTGATATCTTGG 6489
DB 443 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTGATATCTTGG 384
QY 6490 ACTAGACGAGAGTGTGCTGCCACCATCTTGGGGATCTAGATATCAATCAATCACTT 6549
DB 383 AGTAGACGAGAGTGTGCTGCCACCATCTTGGGGATCTAGATATCAATCAATCACTT 332
```

```
QY 6550 GCTTTTGAAGACGTGGTGGAAACGTCTTCTTTTCCACGATGTTCTCGTGGGTGGGGTTC 6609
DB 331 GCTTTTGAAGACGTGGTGGAAACGTCTTCTTTTCCACGATGTTCTCGTGGGTGGGGTTC 272
QY 6610 CATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGACCTTTTCTTTATCGCAAT 6669
DB 271 CATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGACCTTTTCTTTATCGCAAT 212
QY 6670 GATGCAATTTGTAAGCACTTCTTCTTTTCTACTGTCTTTCGATGAAGTACAGATAG 6729
DB 211 GATGCAATTTGTAAGCACTTCTTCTTTTCTACTGTCTTTCGATGAAGTACAGATAG 152
QY 6730 CTGGCAATGGAATCCGAGGAGTTTCCGATATATACCTTTTGTGAAAAGTCTCAATAG 6789
DB 151 CTGGCAATGGAATCCGAGGAGTTTCCGATATATACCTTTTGTGAAAAGTCTCAATAG 92
QY 6790 CCTCTCGTCTTCTGAGACTGTATCTTGAATATCTTGGAGTAGACGAGTGTCTGTCT 6849
DB 91 CCTCTCGTCTTCTGAGACTGTATCTTGAATATCTTGGAGTAGACGAGTGTCTGTCT 32
QY 6850 CCACCATGTTG 6860
DB 31 CCACCATGTTG 21
RESULT 8
AY737283/c
LOCUS       AY737283             12607 bp    DNA             circular SYN 02-MAR-2005
DEFINITION   C-terminal TAP T-DNA vector pYL436, complete sequence.
ACCESSION   AY737283
VERSION     AY737283.1   GI:52630860
KEYWORDS    'C-terminal TAP T-DNA vector pYL436'
SOURCE      C-terminal TAP T-DNA vector pYL436
ORGANISM    C-terminal TAP T-DNA vector pYL436
REFERENCE   1 (bases 1 to 12607)
AUTHORS     Rubio,V., Shen,Y., Saijo,Y., Liu,Y., Gusmaroli,G.,
            Dinesh-Kumar,S.P. and Deng,X.W.
TITLE       An alternative tandem affinity purification strategy applied to
            Arabidopsis protein complex isolation
JOURNAL     Plant J. 41 (5), 767-778 (2005)
PUBMED     15703063
REFERENCE   2 (bases 1 to 12607)
AUTHORS     Liu,Y. and Dinesh-Kumar,S.P.
TITLE       Direct Submission
JOURNAL     Submitted (28-AUG-2004) MCDB, Yale University, New Haven, CT 06520,
            USA
FEATURES             Location/Qualifiers
     1..12607
     /organism="C-terminal TAP T-DNA vector pYL436"
     /mol_type="other DNA"
     /db_xref="taxon:293967"
     /note="derived from pPZP222"
     promoter
     1..769
     /note="duplicated CamV 35S promoter"
     misc_feature
     780..844
     /note="TMV UI omega sequence"
     misc_feature
     845..2718
     /note="attR1-CmrR-cdb-attR2 GATEWAY cassette"
     misc_feature
     2752..3009
     /note="9x myc tag"
     misc_feature
     3163..3180
     /note="6x HIS tag"
     misc_feature
     3181..3204
     /note="protease 3C cleavage site"
     misc_feature
     3244..3591
     /note="2x IgG binding domain"
     misc_feature
     3670..3932
     /note="NOS terminator"
```

ORIGIN		ORIGIN	
Query Match Best Local Similarity Matches 660; Conservative	8.4%; Score 640.2; DB 11; Length 12607; Pred. No. 3.6e-84; 0; Mismatches 3; Indels 8; Gaps 1;	Query Match Best Local Similarity Matches 653; Conservative	8.2%; Score 621; DB 6; Length 5033; Pred. No. 3e-81; 0; Mismatches 30; Indels 4; Gaps 1;
QY	6190 AGATATCACATCAATCCACTTCTTTGAAGAGCGTGGTGGAAACGTCCTTTTCCACGA 6249	QY	6868 AGATATCACATCAATCCACTTCTTTGAAGAGCGTGGTGGAAACGTCCTTTTCCACGA 6927
DB	683 AGATATCACATCAATCCACTTCTTTGAAGAGCGTGGTGGAAACGTCCTTTTCCACGA 624	DB	1103 AGATATCACATCAATCCACTTCTTTGAAGAGCGTGGTGGAAACGTCCTTTTCCACGA 1044
QY	6250 TGTTCTCTGTTGGGTGGGTCCATCTTTGGGACCACTGTTCGGTAGAGCACTCTTGAACGA 6309	QY	6928 TGTTCTCTGTTGGGTGGGTCCATCTTTGGGACCACTGTTCGGTAGAGCACTCTTGAACGA 6987
DB	623 TGTTCTCTGTTGGGTGGGTCCATCTTTGGGACCACTGTTCGGTAGAGCACTCTTGAACGA 564	DB	1043 TGCTCTCTGTTGGGTGGGTCCATCTTTGGGACCACTGTTCGGTAGAGCACTCTTGAACGA 984
QY	6310 TAGCCCTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCC 6369	QY	6988 TAGCCCTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCC 7047
DB	563 TAGCCCTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCC 504	DB	983 TAGCCCTTCTTTATCGCAATGATGGCAATTTGTAGGTGCCACCTTCTCTTTCTACTGTCC 924
QY	6370 TTTGCATGAAGTGACAGATAGCTGGCAATGGAATCCGAGGAGTTCCTCCGATATTACCC 6429	QY	7048 TTTGCATGAAGTGACAGATAGCTGGCAATGGAATCCGAGGAGTTCCTCCGATATTACCC 7107
DB	503 TTTGCATGAAGTGACAGATAGCTGGCAATGGAATCCGAGGAGTTCCTCCGATATTACCC 444	DB	923 TTTGCATGAAGTGACAGATAGCTGGCAATGGAATCCGAGGAGTTCCTCCGATATTACCC 864
QY	6430 TTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTGG 6489	QY	7108 TTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTGG 7167
DB	443 TTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTGG 384	DB	863 TTTGTTGAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTGG 804
QY	6490 AGTAGACGAGAGTGTCTCCACCAGTTGGGATCTAGATATCAATCAATCACTT 6549	QY	7168 AGTAGACGAGAGTGTCTCCACCAGTTGGGATCT---AGATATCAATCAATCACTT 7223
DB	383 AGTAGACGAGAGTGTCTCCACCAGTTG-----ACCATCAATCAATCACTT 332	DB	803 AGTAGACGAGAGTGTCTCCACCAGTTGTGACGGATTTCTAGAGCATCAATCAATCC 744
QY	6550 GCTTTGAACAGCTGGTTGAAAGTCTTCTTTTCCACGATGTCCTCGTGGGTGGGTC 6609	QY	7224 ACTTGTCTTTGAAGAGCTGGTTGGAACGTCCTTTTCCACGATGTCCTCGTGGGTGGG 7283
DB	331 GCTTTGAACAGCTGGTTGAAAGTCTTCTTTTCCACGATGTCCTCGTGGGTGGGTC 272	DB	743 ACTTGTCTTTGAAGAGCTGGTTGGAACGTCCTTTTCCACGATGTCCTCGTGGGTGGG 684
QY	6610 CATCTTTGGGACCACTGTCGCTAGAGGCACTTTGAACAGTATGCTTTCTTTATCGCAAT 6669	QY	7284 GGTCCATCTTTGGGACCACTGTCGCTAGAGGCACTTTGAACAGTATGCTTTCTTTATCG 7343
DB	271 CATCTTTGGGACCACTGTCGCTAGAGGCACTTTGAACAGTATGCTTTCTTTATCGCAAT 212	DB	683 GGTCCATCTTTGGGACCACTGTCGCTAGAGGCACTTTGAACAGTATGCTTTCTTTATCG 624
QY	6670 GATGCAATTTGAGAGCCATCTTCTTTTCTACTGTCTTTGATGTAAGTACGATAG 6729	QY	7344 CAATGATGGCATTTGTAGAGCCATCTTCTTTCTACTGTCTTTGATGTAAGTACGATAG 7403
DB	211 GATGCAATTTGAGAGCCATCTTCTTTTCTACTGTCTTTGATGTAAGTACGATAG 152	DB	623 CAATGATGGCATTTGTAGAGCCATCTTCTTTCTACTGTCTTTGATGTAAGTACGATAG 564
QY	6730 CTGGGCAATGGAATCCGAGAGGTTCCCGATATTACCTTTGTCGAAAGTCTCAATAG 6789	QY	7404 ATAGCTGGGCAATGGAATCCGAGAGGTTCCCGATATTACCTTTGTCGAAAGTCTCA 7463
DB	151 CTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCTTTGTCGAAAGTCTCAATAG 92	DB	563 ATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCTTTGTCGAAAGTCTCA 504
QY	6790 CCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGTCTGTCT 6849	QY	7464 ATAGCCCTCTGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGTCTG 7523
DB	91 CCTCTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGTCTGTCT 32	DB	503 ATAGCCCTTTGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGTGTCTG 444
QY	6850 CCACCATGTTG 6860	QY	7524 TGCTCCACCATGTTGGGATCCACTAG 7550
DB	31 CCACCATGTTG 21	DB	443 TGCTCCACCATGTTGACGGATTTCTAG 417
RESULT 9		RESULT 10	
LOCUS I28266/c		LOCUS CVU10488	
DEFINITION Sequence 1 from patent US 5569828.		DEFINITION Binary cloning vector pPZP121 for plant transformation, complete sequence.	
ACCESSION I28266.1		ACCESSION U10488	
VERSION I28266.1		VERSION U10488.1	
KEYWORDS GI:1919042		KEYWORDS GI:506683	
SOURCE Unknown.		SOURCE Cloning vector pPZP121	
ORGANISM Unknown.		ORGANISM Cloning vector pPZP121	
REFERENCE 1 (bases 1 to 5033)		REFERENCE 1 (bases 1 to 8626)	
AUTHORS McMullen,M.D.; Koch,B.A. and Townsend,R.		AUTHORS Hajdukiewicz,P., Svab,Z. and Maliga,P.	
TITLE Maize chlorotic dwarf virus and resistance thereto		TITLE The small, versatile pp2p family of Agrobacterium binary vectors	
JOURNAL Patent: US 5569828-A 1 29-OCT-1996;		JOURNAL Plant Mol. Biol. 25 (6), 989-994 (1994)	
FEATURES Location/Qualifiers		PUBMED 7919218	
1..5033		REFERENCE 2 (bases 1 to 8626)	
/organism="unknown"			
/mol_type="unassigned DNA"			

AUTHORS Hajdukiewicz, P.  
TITLE Direct Submission  
JOURNAL Submitted (09-JUN-1994) Peter Hajdukiewicz, Waksman Institute,  
Rutgers University, Old Hoes Lane, Piscataway, NJ 08855, USA  
FEATURES  
source  
1. 8626  
/organism="Cloning vector pPZP121"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:35364"  
ORIGIN  
Query Match 8.1%; Score 617.2; DB 11; Length 8636;  
Best Local Similarity 93.6%; Pred. No. 9.2e-81;  
Matches 659; Conservative 0; Mismatches 33; Indels 12; Gaps 1;  
QY 6868 AGATATCACATCAATCCACTTCTTTGAAGAGCTGTTGGAACGCTCTCTTTTCCACGA 6927  
DB 7312 AGATATCACATCAATCCACTTCTTTGAAGAGCTGTTGGAACGCTCTCTTTTCCACGA 7371  
QY 6928 TGTTCCTCGTGGTGGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6987  
DB 7372 TGTTCCTCGTGGTGGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 7431  
QY 6988 TAGCCTTTCCTTATCGCAATGATGGCAATTTGAGGCTGCTTCTTCTTCTCTGTC 7047  
DB 7432 TAGCCTTTCCTTATCGCAATGATGGCAATTTGAGGCTGCTTCTTCTTCTCTGTC 7491  
QY 7048 TTTTCATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 7107  
DB 7492 TTTTCATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 7551  
QY 7108 TTTTGTGAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTCTGG 7167  
DB 7552 TTTTGTGAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTCTGG 7611  
QY 7168 AGTAGACGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTT 7227  
DB 7612 AGTAGACGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTT 7659  
QY 7228 GCTTTGAAGACGTGGTTGAAAGCTCTTTTTCACAGATGTTCTCGTGGTGGGGTTC 7287  
DB 7660 GCTTTGAAGACGTGGTTGAAAGCTCTTTTTCACAGATGTTCTCGTGGTGGGGTTC 7719  
QY 7288 CATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCTTTATCGCAAT 7347  
DB 7720 CATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCTTTATCGCAAT 7779  
QY 7348 GATGGCAATTTGTAAGCCATCTTCTTTCTACTGCTCTTCTGATGAAGTGACAGATAG 7407  
DB 7780 GATGGCAATTTGTAAGCCATCTTCTTTCTACTGCTCTTCTGATGAAGTGACAGATAG 7839  
QY 7408 CTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCTTTGTTGAAAAGTCTCAATAG 7467  
DB 7840 CTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCTTTGTTGAAAAGTCTCAATAG 7899  
QY 7468 CCTCTGTGCTCTTCTGAGACTGTATCTTTGATATTCTTTGAGTAGAGAGTGTCTGTCT 7527  
DB 7900 CCTTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTTGAGTAGAGAGTGTCTGTCT 7959  
QY 7528 CCACCATGTTGGGATCCACTAGTTCTTAGAGCGCGCCGCCACCGC 7571  
DB 7960 CCACCATGTTGGGATCCACTAGTTCTTAGCAATAGCGAAACCGCCTC 8003  
RESULT 11  
CVU10491  
LOCUS 8731 bp DNA circular SYN 28-JAN-1995  
DEFINITION Binary cloning vector pPZP221 for plant transformation, complete  
sequence.  
ACCESSION U10491  
VERSION U10491.1 GI:506686  
KEYWORDS  
SOURCE Cloning vector pPZP221

ORGANISM Cloning vector pPZP221  
other sequences; artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 8731)  
AUTHORS Hajdukiewicz, P., Svab, Z. and Maliga, P.  
TITLE The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation  
JOURNAL Plant Mol. Biol. 25 (6), 989-994 (1994)  
PUBMED 7919218  
REFERENCE 2 (bases 1 to 8731)  
AUTHORS Hajdukiewicz, P.  
TITLE Direct Submission  
JOURNAL Submitted (09-JUN-1994) Peter Hajdukiewicz, Waksman Institute,  
Rutgers University, Old Hoes Lane, Piscataway, NJ 08855, USA  
FEATURES  
source  
1. 8731  
/organism="Cloning vector pPZP221"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:35371"  
ORIGIN  
Query Match 8.1%; Score 617.2; DB 11; Length 8731;  
Best Local Similarity 93.6%; Pred. No. 9.1e-81;  
Matches 659; Conservative 0; Mismatches 33; Indels 12; Gaps 1;  
QY 6868 AGATATCACATCAATCCACTTCTTTGAAGAGCTGTTGGAACGCTCTCTTTTCCACGA 6927  
DB 7417 AGATATCACATCAATCCACTTCTTTGAAGAGCTGTTGGAACGCTCTCTTTTCCACGA 7476  
QY 6928 TGTTCCTCGTGGTGGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6987  
DB 7477 TGTTCCTCGTGGTGGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 7536  
QY 6988 TAGCCTTTCCTTATCGCAATGATGGCAATTTGAGGCTGCTTCTTCTTCTGTC 7047  
DB 7537 TAGCCTTTCCTTATCGCAATGATGGCAATTTGAGGCTGCTTCTTCTTCTGTC 7596  
QY 7048 TTTTCATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 7107  
DB 7597 TTTTCATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 7656  
QY 7108 TTTTGTGAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTCTGG 7167  
DB 7657 TTTTGTGAAAGTCTCAATAGCCCTTGGTCTTCTGAGACTGTATCTTTGATATTCTGG 7716  
QY 7168 AGTAGACGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTT 7227  
DB 7717 AGTAGACGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTT 7764  
QY 7228 GCTTTGAAGACGTGGTTGGAACGCTCTTTTTCACAGATGTTCTCGTGGTGGGGTTC 7287  
DB 7765 GCTTTGAAGACGTGGTTGGAACGCTCTTTTTCACAGATGTTCTCGTGGTGGGGTTC 7824  
QY 7288 CATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCTTTATCGCAAT 7347  
DB 7825 CATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCTTTATCGCAAT 7884  
QY 7348 GATGGCAATTTGTAAGCCATCTTCTTTCTACTGCTCTTCTGATGAAGTGACAGATAG 7407  
DB 7885 GATGGCAATTTGTAAGCCATCTTCTTTCTACTGCTCTTCTGATGAAGTGACAGATAG 7944  
QY 7408 CTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCTTTGTTGAAAAGTCTCAATAG 7467  
DB 7945 CTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCTTTGTTGAAAAGTCTCAATAG 8004  
QY 7468 CCTCTGTGCTCTGAGACTGTATCTTTGATATTCTTTGAGTAGAGAGTGTCTGTCT 7527  
DB 8005 CCTTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTTGAGTAGAGAGTGTCTGTCT 8064  
QY 7528 CCACCATGTTGGGATCCACTAGTTCTTAGAGCGCGCCGCCACCGC 7571  
DB 8065 CCACCATGTTGGGATCCACTAGTTCTTAGCAATAGCGAAACCGCCTC 8108



```

DEFINITION Binary vector pCambia-2200, complete sequence.
ACCESSION AF234313.1 GI:7638136
VERSION
SOURCE
ORGANISM
REFERENCE
AUTHORS Hajdukiewicz, P., Svab, Z. and Maliga, P.
TITLE The small, versatile pP2P family of Agrobacterium binary vectors
for plant transformation
JOURNAL Plant Mol. Biol. 25 (6), 989-994 (1994)
PUBMED 7919218
REFERENCE
AUTHORS 2 (bases 1 to 8894)
Roberts, C., Rajagopal, S., Smith, L.M., Nguyen, T.A., Yang, W.,
Nugrohu, S., Ravi, K.S., Vijayachandra, K., Harcourt, R.L.,
Dransfield, L., Desamero, N., Slamet, I., Hadjukiewicz, P., Svab, Z.,
Maliga, P., Mayer, J.E., Keese, P.K., Kilian, A. and Jefferson, R.A.
A comprehensive set of modular vectors for advanced manipulations
and efficient transformation of plants
Unpublished
Full description of constructs
3 (bases 1 to 8894)
Roberts, C., Rajagopal, S., Smith, L.M., Nguyen, T.A., Yang, W.,
Nugrohu, S., Ravi, K.S., Vijayachandra, K., Harcourt, R.L.,
Dransfield, L., Desamero, N., Slamet, I., Hadjukiewicz, P., Svab, Z.,
Maliga, P., Mayer, J.E., Keese, P.K., Kilian, A. and Jefferson, R.A.
Direct Submission
Submitted (15-FEB-2000) CAMBIA, Clunies Ross St, Black Mountain /
GPO Box 3200, Canberra, ACT 2601, Australia
FEATURES
source
Location/Qualifiers
1..8894
/organism="Binary vector pCambia-2200"
/mol_type="other DNA"
/db_xref="taxon:118405"
complement(980..1980)
/notes="STA region from pVSI plasmid"
complement(2573..3573)
/notes="pVSI-REP; replication origin from pVSI"
complement(3983..4243)
/notes="bom site from pBR322"
complement(4383..4663)
/notes="pBR322 origin of replication"
complement(5138..5797)
/notes="cat (chloramphenicol resistance) gene from pBR325"
/codon_start=1
/product="chloramphenicol acetyl transferase"
/protein_id="AAF65392.1"
/db_xref="GI:7638138"
/translation="MEKKTGYTTVDISQHRKEHFEAFQSVQAQCTYNQTVQLDITAF
LKTVKXKHFPYPAFHILARLMAHPEPRMAKDGLVWDSVHPCTVPHEQETEF
SSLSEYHDDPQFHLIYSQDVACGENLAYPKGFLENPFVSNPWPVSTFDLNV
ANWDDFFAPVFTMGKYTGDKVLMPLAQVHVHNCDFHVGEMLNELQYCDENQGG
A"
misc_feature 6325..6350
/notes="left border repeat from C58 T-DNA"
misc_feature 6417..6620
/notes="CamV 3'UTR (polyA signal)"
CDS complement(6658..7455)
/notes="aadA (kanamycin resistance) gene amplified from
pIG121Hm"
/codon_start=1
/product="aminoglycoside phosphotransferase"
/protein_id="AAF65391.1"
/db_xref="GI:7638137"
/translation="MGIEQDGLHAGSPAAWVERLFQYDWAQQTIGSDIAVPRLSAQG
RPVLFYKTLQSGALNELQDEARLTLWLTATGVPCEAVLDVTEAGRDWLLGLVFGQD
LLSLHAPAKSYIMADAMRLRLTDPATCFDHOAKHRIERARTMEAGLVQDDLD
EBHQGLAPAEFLARLARPMDGDELVVTHGDACLPMIMVNGRFGSFDICGRIGVADR
YQDIALATRDIAELGGEWADRFLVLYGIAAPDSQRIAPRYRLDREFF"
complement(7486..8293)
/promoter /notes="CamV35S; 35S promoter from CamV"
8414..8468
/promoter

```

```

/notes="lacZ; lacZ promoter"
8496..8729
/codon_start=1
/product="LacZ alpha fragment"
/protein_id="AAF65393.1"
/db_xref="GI:7638139"
/translation="MTMLTNSSSVPGDPLESTCRHSLASLALAVVLQRDWNENPGVTQLN
RLAHPFPASWRNSEATDRPSQQLRSINGEC"
misc_feature 8510..8566
/notes="pUC18 MCS; polylinker"
misc_feature 8808..8833
/notes="right border T-DNA repeat"
ORIGIN
Query Match 8.1%; Score 617.2; DB 11; Length 8894;
Best Local Similarity 93.6%; Pred. No. 9.1e-81;
Matches 659; Conservative 0; Mismatches 33; Indels 12; Gaps 1;
QY 6868 AGATATACATCAATCCACTTCTTGAAGACGTGTGGAACGTCCTCTTTTCCACGA 6927
DB 7612 AGATATACATCAATCCACTTCTTGAAGACGTGTGGAACGTCCTCTTTTCCACGA 7671
QY 6928 TCTTCTCTGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6987
DB 7672 TCTTCTCTGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 7731
QY 6988 TAGCCTTCTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTTTTCTACTGTCC 7047
DB 7732 TAGCCTTCTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTTTTCTACTGTCC 7791
QY 7048 TTTTCATGAGTACAGATAGTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 7107
DB 7792 TTTTCATGAGTACAGATAGTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 7851
QY 7108 TTTGTTGAAAAGTCTCAATAGCCCTCTGTCTTCTGAGACTGTATCTTTGATATCTTGG 7167
DB 7852 TTTGTTGAAAAGTCTCAATAGCCCTTGTCTTCTGAGACTGTATCTTTGATATCTTGG 7911
QY 7168 AGTAGACGAGAGTGTGGTCTCCACATGTTGGGATCTAGATATACATCAATCCACTT 7227
DB 7912 AGTAGACGAGAGTGTGGTCTCCACATGTTGGGATCTAGATATACATCAATCCACTT 7959
QY 7228 GCTTTGAGACGTGTGGAGCGTCTTCTTTTCCACGATGTTCCTCGGGTGGGGGTC 7287
DB 7960 GCTTTGAGACGTGTGGAGCGTCTTCTTTTCCACGATGTTCCTCGGGTGGGGGTC 8019
QY 7288 CATCTTTGGGACCACTGTCGGTAGAGGCATCTTTGAACGATAGCCTTTCTTTATCGCAAT 7347
DB 8020 CATCTTTGGGACCACTGTCGGTAGAGGCATCTTTGAACGATAGCCTTTCTTTATCGCAAT 8079
QY 7348 GATGCAATTTGAGAACCATCTTCTTTTCTACTGTCTTTCGATGAAAGTGACAGATAG 7407
DB 8080 GATGCAATTTGAGGTGCCACCTTCTCTTTTCTACTGTCTTTCGATGAAAGTGACAGATAG 8139
QY 7408 CTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCTTTGTTGAAAAGTCTCAATAG 7467
DB 8140 CTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCTTTGTTGAAAAGTCTCAATAG 8199
QY 7468 CCTCTGTGCTTCTGAGACTGTATCTTTGATATTTCTTGAGTAGACGAGTGTCTGCT 7527
DB 8200 CCTTTGCTTCTGAGACTGTATCTTTGATATTTCTTGAGTAGACGAGTGTCTGCT 8259
QY 7528 CCACCATGTTGGGATGCCACTAGTTCTTAGAGCGCGCCGCCCGC 7571
DB 8260 CCACCATGTTGGCAAGTCTCTAGCCAAATAGCAAAACCGCTC 8303
RESULT 14
CVU10487
LOCUS
DEFINITION Binary cloning vector pP2P111 for plant transformation, complete
sequence.
CVU10487
ACCESSION U10487

```



VERSION	U10487.1	GI:506682	Db	8243	CCACCATGTTGGCAAGCTGCTCTAGCCAATAGCAACCGCTC	8286
KEYWORDS	Cloning vector pPZp111					RESULT 15
SOURCE	other sequences; artificial sequences; vectors.					AF234296
ORGANISM	1 (bases 1 to 8909)					AF234296
REFERENCE	Hajdukiewicz, P., Svab, Z. and Maliga, P.					AF234296
AUTHORS	The small, versatile pPZP family of Agrobacterium binary vectors					AF234296.1
TITLE	for plant transformation					GI:7638064
JOURNAL	Plant Mol. Biol. 25 (6), 989-994 (1994)					
PUBMED	7919218					
REFERENCE	2 (bases 1 to 8909)					
AUTHORS	Hajdukiewicz, P.					
JOURNAL	Direct Submission					
TITLE	Submitted (09-JUN-1994) Peter Hajdukiewicz, Waksman Institute,					
JOURNAL	Rutgers University, Old Hoes Lane, Piscataway, NJ 08855, USA					
FEATURES	Location/Qualifiers					
source	1..8909					
	/organism="Cloning vector pPZp111"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:35362"					
ORIGIN						
Query Match	8.1%; Score 617.2; DB 11; Length 8909;					
Best Local Similarity	93.6%; Pred. No. 9.1e-81;					
Matches	659; Conservative 0; Mismatches 33; Indels 12; Gaps 1;					
QY	6868	AGATATCACATCAATCCACTGCTTTGAAGACGTGGTTGGAACTGCTCTCTTTTCCACGA	6927			
Db	7595	AGATATCACATCAATCCACTGCTTTGAAGACGTGGTTGGAACTGCTCTCTTTTCCACGA	7654			
QY	6928	TGTTCTCTGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGATCTTGAACGA	6987			
Db	7655	TGTTCTCTGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGATCTTGAACGA	7714			
QY	6988	TAGCCTTTCTTTATCGCAATGATGCACTTTCTAGAACCACTCTCTTTCTACTGTCC	7047			
Db	7715	TAGCCTTTCTTTATCGCAATGATGCACTTTCTAGAACCACTCTCTTTCTACTGTCC	7774			
QY	7048	TTTCAGTAAGTACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC	7107			
Db	7775	TTTCAGTAAGTACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC	7834			
QY	7108	TTTGTGAAAGTCTCAATAGCCCTCTGTTCTTGAGACTGTATCTTTGATATCTTGG	7167			
Db	7835	TTTGTGAAAGTCTCAATAGCCCTCTGTTCTTGAGACTGTATCTTTGATATCTTGG	7894			
QY	7168	AGTAGACGAGTGTGCTGCTCCACCATGTTGGGATCTAGATATCACATCAATCCACTT	7227			
Db	7895	AGTAGACGAGTGTGCTGCTCCACCATGTTGGGATCTAGATATCACATCAATCCACTT	7942			
QY	7228	GCTTTGAAGACGTGGTTGAAACGTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTTC	7287			
Db	7943	GCTTTGAAGACGTGGTTGAAACGTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTTC	8002			
QY	7288	CATCTTTGGGACCACTGTGCGTAGAGGCACTTTGAAACGATAGCTTTTATCGCAAT	7347			
Db	8003	CATCTTTGGGACCACTGTGCGTAGAGGCACTTTGAAACGATAGCTTTTATCGCAAT	8062			
QY	7348	GATGGCAATTTGTAGAGCCATCTTCTCTTCTACTGCTCTTTCGATGAAGTACAGATAG	7407			
Db	8063	GATGGCAATTTGTAGAGTGCACCTCTCTTCTACTGCTCTTTCGATGAAGTACAGATAG	8122			
QY	7408	CTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCTTTGTTGAAAGTCTCAATAG	7467			
Db	8123	CTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCTTTGTTGAAAGTCTCAATAG	8182			
QY	7468	CCCTCTGCTTCTGAGACTGTATCTTTGATATCTTTGAGTAGACGAGTGTCTGTCT	7527			
Db	8183	CCCTCTGCTTCTGAGACTGTATCTTTGATATCTTTGAGTAGACGAGTGTCTGTCT	8242			
QY	7528	CCACCATGTTGGGATCCCACTAGTTCTAGAGCGCGCCACCGC	7571			



Search completed: December 31, 2005, 00:12:40  
Job time : 24349 secs

[illegible]

**BEST AVAILABLE COPY**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 31, 2005, 00:12:49 ; Search time 2698 Seconds  
(without alignments)  
18724.377 Million cell updates/sec

Title: US-10-650-249-1  
Perfect score: 7580  
Sequence: 1 agctctattattcaagaga.....gcgcaccgcggtgagct 7580

Scoring table: OLIGO\_NUC  
Gapop\_60.0, Gapext 60.0

Searched: 4996997 seqs, 332346308 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_21.\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*  
14: Geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7580	100.0	7580	ADL71829	Adl71829 Arabidops
2	1503	19.8	1746	ADL71846	Adl71846 Arabidops
3	1402	18.5	10078	ABQ73047	Abq73047 Tomato an
4	1367	18.0	17511	ADV39061	Adv39061 Plant gen
5	1367	18.0	18987	ADV39062	Adv39062 Plant gen
6	1360	17.9	1361	AAA88401	Aaa88401 4x CamV 3
7	1235	16.3	1235	ADL71845	Adl71845 Arabidops
8	1062	14.0	1062	ADL71848	Adl71848 Arabidops
9	1058	14.0	1058	ADL71847	Adl71847 Arabidops
10	888	11.7	888	ADL71840	Adl71840 Arabidops
11	684	9.0	684	ADL71843	Adl71843 Arabidops
12	581	7.7	645	ADK59804	Adk59804 Plant DNA
13	581	7.7	795	ADD30377	Add30377 Plant yie
14	581	7.7	795	ADL71840	Adl71840 Arabidops
15	577	7.6	577	ADL71847	Adl71847 Arabidops
16	368	4.9	2000	ABX56844	Abx56844 Arabidops
17	331	4.4	12739	ADW39084	Adw39084 Binary ve
18	300	4.0	552	AAD39450	Aad39450 FMV35S46
19	299	3.9	323	AAC87191	Aac87191 CamV35S p

ALIGNMENTS

RESULT 1

ADL71829

ID ADL71829 standard; DNA; 7580 BP.

XX ADL71829;

XX 20-MAY-2004 (first entry)

XX Arabidopsis thaliana OBP3 DNA.

XX Transgenic plant; Dof transcription factor; ocs binding factor;  
KW plant size; plant stature; root growth; plant; gene; db; OBF;  
KW OBF binding protein; OBF3; SOB1; mouse-ear cross.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers

FT CDS 4103..5119

FT /\*tag= a

FT /product= "Arabidopsis thaliana OBP3 protein"

FT /transl\_except= (pos:4811..4909, aa:Gly-Gly)

XX US2004045055-A1.

XX 04-MAR-2004.

XX 28-AUG-2003; 2003US-00650249.

XX 28-AUG-2002; 2002US-0406657P.

(UNIW ) UNIV WASHINGTON.

XX Neff MM;

XX WPI; 2004-225757/21.

XX P-PSDB; ADL71830.

XX New transgenic plant cell, useful in producing plants with altered size  
and stature and with normal and healthy root growth.

XX Claim 24; SEQ ID NO 1; 53pp; English.

XX

CC	The invention relates to a transgenic plant transformed by a Dof														
CC	transcription factor, OBF (ocs binding factor) binding protein (OBP3).														
CC	OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic														
CC	acid and polypeptides are useful in producing transgenic plants with														
CC	altered size and stature and with normal and healthy root growth. The														
CC	present sequence is Arabidopsis thaliana OBP3 DNA.														
XX															
SQ	Sequence 7580 BP; 2376 A; 1376 C; 1291 G; 2537 T; 0 U; 0 Other;														
	Query Match	100.0%;	Score 7580;	DB 12;	Length 7580;										
	Best Local Similarity	100.0%;	Pred. No. 0;												
	Matches 7580;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;										
QY	1	AGCTCTATTAAATTCAGAGAGAGCAGCAATAAAGCGAAAAA	CTCAAAACCTAAGTTTCTCGA	60											
DB	1	AGCTCTATTAAATTCAGAGAGAGCAGCAATAAAGCGAAAAA	CTCAAAACCTAAGTTTCTCGA	60											
QY	61	ATATGAAGGGTTAGATAATCAATTCCTCAACTAGTTAAAAA	AGTAATGATAAAATTA	120											
DB	61	ATATGAAGGGTTAGATAATCAATTCCTCAACTAGTTAAAAA	AGTAATGATAAAATTA	120											
QY	121	AAAAACAAATGATCAATTAAGAGACAGTAGTTATGATATAT	ATATGTTGGGATCGAATTA	180											
DB	121	AAAAACAAATGATCAATTAAGAGACAGTAGTTATGATATAT	ATATGTTGGGATCGAATTA	180											
QY	181	GTTGACATCAAAAGATCAAAATAATGACCGGTAAATGACC	CGGTAAATCCAAAGCGCAAT	240											
DB	181	GTTGACATCAAAAGATCAAAATAATGACCGGTAAATGACC	CGGTAAATCCAAAGCGCAAT	240											
QY	241	CATGCGATATCAAAACACCTTAATGTCTCAATCTCAATTA	CTCAATTA	300											
DB	241	CATGCGATATCAAAACACCTTAATGTCTCAATCTCAATTA	CTCAATTA	300											
QY	301	CAAAAGCTAGCTAGATTTTCTCAAGTGAGCAACAGTCTAA	TTCTTCTGAAAAA	360											
DB	301	CAAAAGCTAGCTAGATTTTCTCAAGTGAGCAACAGTCTAA	TTCTTCTGAAAAA	360											
QY	361	TTTTCCGAGTGTAATATCCAAATCTACTTACAAATTTGAG	CATTAATCTAGTTTCT	420											
DB	361	TTTTCCGAGTGTAATATCCAAATCTACTTACAAATTTGAG	CATTAATCTAGTTTCT	420											
QY	421	CTGCAACTTTAAACACAGGTGTACAAGTGTCAACACAGAT	CTAGCGTAAACACTTAAG	480											
DB	421	CTGCAACTTTAAACACAGGTGTACAAGTGTCAACACAGAT	CTAGCGTAAACACTTAAG	480											
QY	481	CTAGTACTTTAAATAGATTTATGCCCTATTTTTCGACCAT	TATATATATAATTTCCAGCCT	540											
DB	481	CTAGTACTTTAAATAGATTTATGCCCTATTTTTCGACCAT	TATATATATAATTTCCAGCCT	540											
QY	541	TTTCGTGAAAAAATGCGCATGTTCTTGTTGGAATCTCTCT	CTCTACTAAGATTG	600											
DB	541	TTTCGTGAAAAAATGCGCATGTTCTTGTTGGAATCTCTCT	CTCTACTAAGATTG	600											
QY	601	GCATGCACTGGTAAACGATTTCCATTTGTATATCTATCTCG	ATATTTTCCACCTTAAATAT	660											
DB	601	GCATGCACTGGTAAACGATTTCCATTTGTATATCTATCTCG	ATATTTTCCACCTTAAATAT	660											
QY	661	CTTGAAAAATTAAGATCAAAATATATGAGAACATATATTT	GTTGTTGTAATAGCCT	720											
DB	661	CTTGAAAAATTAAGATCAAAATATATGAGAACATATATTT	GTTGTTGTAATAGCCT	720											
QY	721	TATATAGTGTGATGATGTAATGAGAACATATATGTA	CTATGTTGTTGTAATAGCCT	780											
DB	721	TATATAGTGTGATGATGTAATGAGAACATATATGTA	CTATGTTGTTGTAATAGCCT	780											
QY	781	AGAAGGTGTTGCTTTTACAACTATGAATATGGGCTCAT	TGAGATGTACACTACAGATAAGCG	840											
DB	781	AGAAGGTGTTGCTTTTACAACTATGAATATGGGCTCAT	TGAGATGTACACTACAGATAAGCG	840											
QY	841	AAGATTCCTGTTGTCATGAAATGTTGTTACTAATAA	AAAAAACAATGCAATTTCTATA	900											
DB	841	AAGATTCCTGTTGTCATGAAATGTTGTTACTAATAA	AAAAAACAATGCAATTTCTATA	900											

Qy	901	AGACAAATTTCTTTAAATAATAAATTTCTATAA	CAAAATAAGATAAGTGTCTCTTTAAA	960
Db	901	AGACAAATTTCTTTAAATAATAAATTTCTATAA	CAAAATAAGATAAGTGTCTCTTTAAA	960
Qy	961	AACATGCAAAAGAATAATATAGATTACCGTAT	CGATTCAGATTTTCATACAAATTTTATAT	1020
Db	961	AACATGCAAAAGAATAATATAGATTTTACCGTAT	CGATTTTCATACAAATTTTATATATTTT	1020
Qy	1021	TGAGCTTGAAGATTAAACATGACAAACTGTAT	CGTGTCTCTGCTATTTCAACCCCTAGA	1080
Db	1021	TGAGCTTGAAGATTAAACATGACAAACTGTAT	CGTGTCTCTGCTATTTCAACCCCTAGA	1080
Qy	1081	AGAAAGTGAACATGGAACCTTTATGTATTTG	ATGATACGGCGAGCTAGTCTTCTCCCTAC	1140
Db	1081	AGAAAGTGAACATGGAACCTTTATGTATTTG	ATGATACGGCGAGCTAGTCTTCTCCCTAC	1140
Qy	1141	TCCAATAGATGAAGACATTTATCACTCAGGTT	CAGCTACTTCGAAGCGCAACATATCGACA	1200
Db	1141	TCCAATAGATGAAGACATTTATCACTCAGGTT	CAGCTACTTCGAAGCGCAACATATCGACA	1200
Qy	1201	AAAAATCGTTTTAGCTCTATCATCTGTCTTT	TGAAGAAAAATATCAACATATCAAAATACA	1260
Db	1201	AAAAATCGTTTTAGCTCTATCATCTGTCTTT	TGAAGAAAAATATCAACATATCAAAATACA	1260
Qy	1261	TATACACACTCCCAAAATATATAACCAATATA	TATATGTTACTACGAAATTTCCAAT	1320
Db	1261	TATACACACTCCCAAAATATATAACCAATATA	TATATGTTACTACGAAATTTCCAAT	1320
Qy	1321	GATATTTGCTTTTGAACAACACTTAACTGT	GAATTAACAAGCTAAGGCAATATATCTAT	1380
Db	1321	GATATTTGCTTTTGAACAACACTTAACTGT	GAATTAACAAGCTAAGGCAATATATCTAT	1380
Qy	1381	ATTCTTTCTTTGTTGCTCAACTTCTTAAGT	TAATTTCAAAATTTGGAAGTTT	1440
Db	1381	ATTCTTTCTTTGTTGCTCAACTTCTTAAGT	TAATTTCAAAATTTGGAAGTTT	1440
Qy	1441	TGTCATCTAAATTTGAAGTTCTTTTACCGG	ATCATTTTGTCTCAGGTTGATATATCTTT	1500
Db	1441	TGTCATCTAAATTTGAAGTTCTTTTACCGG	ATCATTTTGTCTCAGGTTGATATATCTTT	1500
Qy	1501	CTTAGTCTGATACGAAAACTTTATATAAT	ATATGATAGAGAGACATGTTTGATTTA	1560
Db	1501	CTTAGTCTGATACGAAAACTTTATATAAT	ATATGATAGAGAGACATGTTTGATTTA	1560
Qy	1561	TATTTTCTGTTGTTAAAAACATCTTGATTA	TGAACTATATAGTTAGGGATGTTTGTAT	1620
Db	1561	TATTTTCTGTTGTTAAAAACATCTTGATTA	TGAACTATATAGTTAGGGATGTTTGTAT	1620
Qy	1621	TGTTGTCGACATAGTGAGTCCATTTAAAA	AGAGGCTCTGATTTAAATTTTACGTTT	1680
Db	1621	TGTTGTCGACATAGTGAGTCCATTTAAAA	AGAGGCTCTGATTTAAATTTTACGTTT	1680
Qy	1681	CACAAATCTTTCTTTTAGAAAATCGGGA	CTGGGACACCTTCCCTAACAATGTCCTTT	1740
Db	1681	CACAAATCTTTCTTTTAGAAAATCGGGA	CTGGGACACCTTCCCTAACAATGTCCTTT	1740
Qy	1741	ACTAATCTTACGTACCCCTCACATTCGT	TAACCAATCAATATAATATAGAGAC	1800
Db	1741	ACTAATCTTACGTACCCCTCACATTCGT	TAACCAATCAATATAATATAGAGAC	1800
Qy	1801	TGTTGATCATAAATTCGAAATATTTTT	CACTAATTTCAATGTTATCGGTAAGTTA	1860
Db	1801	TGTTGATCATAAATTCGAAATATTTTT	CACTAATTTCAATGTTATCGGTAAGTTA	1860
Qy	1861	GGTATAACATCAAGAAATCACGAAAGAA	TTTAAACCAACCTTGTGCAATCATGAT	1920
Db	1861	GGTATAACATCAAGAAATCACGAAAGAA	TTTAAACCAACCTTGTGCAATCATGAT	1920
Qy	1921	TTTTGGCTTATACCTTTCTAAATTTTT	TATATCTTGTGCTGCAAAATTAGCACC	1980
Db	1921	TTTTGGCTTATACCTTTCTAAATTTTT	TATATCTTGTGCTGCAAAATTAGCACC	1980
Qy	1981	TATTTCTTCTTCAACATCGAATCTTT	TATTTTGTAAAGGCAATTTTTTCTTAA	2040



QY	4201	TCGAGTTCAGCTCAAGCTAGAGTGAATTCATAGGTGGAACTGCTCGGATCGCAAAAGT	4260	QY	5281	ATCTTTTAAAGATCTTCAAAGTGTGAGTAGTGTATTTATTTGGTTGGCTTCTGGTGATATTTA	5340
Db	4201	TCGAGGTTCGAGTCAAGCTAGAGTGAATTCATAGGTGGAACTGCTCGGATCGCAAAAGT	4260	Db	5281	ATCTTTTAAAGATCTTCAAAGTGTGAGTAGTGTATTTATTTGGTTGGCTTCTGGTGATATTTA	5340
QY	4261	CCCATTTGCCCTGAAGCAGCTCTAAATTTGCCCTAGAGTGTGACTCAACCAATACCTAAAGTTCTG	4320	QY	5341	TGTTTTTATAGAAATTTGGTCTTATATATTTGGTCTATATATAGAGGTGTGGGTGATATGAT	5400
Db	4261	CCCATTTGCCCTGAAGCAGCTCTAAATTTGCCCTAGAGTGTGACTCAACCAATACCTAAAGTTCTG	4320	Db	5341	TGTTTTTATAGAAATTTGGTCTTATATATTTGGTCTATATATAGAGGTGTGGGTGATATGAT	5400
QY	4321	TTACTTCAATAAATAATAGCCTTACTCAACCTCGCCATTTCTGCATAAAACATATGCTGGCTA	4380	QY	5401	GAAATTCAGAGTGTGATGTGGAACTTTTTTGTGTGTTTCATTTGAATAATCATCGAATTCCT	5460
Db	4321	TTACTTCAATAAATAATAGCCTTACTCAACCTCGCCATTTCTGCATAAAACATATGCTGGCTA	4380	Db	5401	GAAATTCAGAGTGTGATGTGGAACTTTTTTGTGTGTTTCATTTGAATAATCATCGAATTCCT	5460
QY	4381	TTGGACAGCTGGCGGTTCTTTCAGGAAATGTTCTCTGTGTGGAGGCTTTAGAGGAACAA	4440	QY	5461	CAATTTCTTGAGAGCCCAATATGAGACATTTGAGACATCTATAGAACATATATGTAATGTA	5520
Db	4381	TTGGACAGCTGGCGGTTCTTTCAGGAAATGTTCTCTGTGTGGAGGCTTTAGAGGAACAA	4440	Db	5461	CAATTTCTTGAGAGCCCAATATGAGACATTTGAGACATCTATAGAACATATATGTAATGTA	5520
QY	4441	GAGAGCAAAATCCAGATCGAAATCTACGGTCTGGTCTCGACTGATATACTACTAGTAC	4500	QY	5521	TATTTAAACGTACTTAAAGTCGAATTTTATGACCAAAAGTAAATAAATATATGCGGAAATGTA	5580
Db	4441	GAGAGCAAAATCCAGATCGAAATCTACGGTCTGGTCTCGACTGATATACTACTAGTAC	4500	Db	5521	TATTTAAACGTACTTAAAGTCGAATTTTATGACCAAAAGTAAATAAATATATGCGGAAATGTA	5580
QY	4501	TTTCAATCACTTACTTCTCGCCCAAGTTTACTCAAAACCTTAGCAAGTTTTCATAGCTACGGTCA	4560	QY	5581	TGCTAATATCGAGTTTAAACTATTTTCCCAATATAACAACATATTTCTCTTTTCGTCCTT	5640
Db	4501	TTTCAATCACTTACTTCTCGCCCAAGTTTACTCAAAACCTTAGCAAGTTTTCATAGCTACGGTCA	4560	Db	5581	TGCTAATATCGAGTTTAAACTATTTTCCCAATATAACAACATATTTCTCTTTTCGTCCTT	5640
QY	4561	AATCCGGAGTTTAAATTCOAATCTGCGCATCTTGCTCTCTCTCCAAAGCCTTGGAGATT	4620	QY	5641	CTTATATATCTTATTTCTGATTTCTTATTTTCTTTTAAATTCCTTTTCTTTTCTTTTCCCAA	5700
Db	4561	AATCCGGAGTTTAAATTCOAATCTGCGCATCTTGCTCTCTCTCCAAAGCCTTGGAGATT	4620	Db	5641	CTTATATATCTTATTTCTGATTTCTTATTTTCTTTTAAATTCCTTTTCTTTTCTTTTCCCAA	5700
QY	4621	CAATTCAGCAACACTGGATTTAGATTTTGTGTGAACTCAAAATAAGCAATGATAGTGG	4680	QY	5701	GACACAAAAAATAAATAACAGAAAAAGAGATTTTAAAAATTTCAATAACCCAC	5760
Db	4621	CAATTCAGCAACACTGGATTTAGATTTTGTGTGAACTCAAAATAAGCAATGATAGTGG	4680	Db	5701	GACACAAAAAATAAATAACAGAAAAAGAGATTTTAAAAATTTCAATAACCCAC	5760
QY	4681	TATGAGTCTTCTAGTGTGGATCTTGGATGCATGGAGAAATACCTCCATCAACAAGCTCA	4740	QY	5761	GAGAAATTTATGCTTCAATCATATGCGCAACTAACTAAATTTGAAAAAGACAATGGAA	5820
Db	4681	TATGAGTCTTCTAGTGTGGATCTTGGATGCATGGAGAAATACCTCCATCAACAAGCTCA	4740	Db	5761	GAGAAATTTATGCTTCAATCATATGCGCAACTAACTAAATTTGAAAAAGACAATGGAA	5820
QY	4741	GCAATTCCTTTCTTGATCAACAATCTACCGGATTTGGTGCAATCTTCAAAACCGCTTATATCC	4800	QY	5821	CGATTTAATATTTGCTTCAATCATATGCGCAACTAACTAAATTTGAAAAAGACAATGGAA	5880
Db	4741	GCAATTCCTTTCTTGATCAACAATCTACCGGATTTGGTGCAATCTTCAAAACCGCTTATATCC	4800	Db	5821	CGATTTAATATTTGCTTCAATCATATGCGCAACTAACTAAATTTGAAAAAGACAATGGAA	5880
QY	4801	ATTACTAGAGGTAAAGGAGGTGTTAATCAAGTGATTTCTCAACAGAGAGTATGATT	4860	QY	5881	TGACTGAAACCAATGATATCTCTCAAGTCTCAACCTATGAAAGAAATCATGTAACCAATAG	5940
Db	4801	ATTACTAGAGGTAAAGGAGGTGTTAATCAAGTGATTTCTCAACAGAGAGTATGATT	4860	Db	5881	TGACTGAAACCAATGATATCTCTCAAGTCTCAACCTATGAAAGAAATCATGTAACCAATAG	5940
QY	4861	TTCCCAATCAGCTAAATGTTTAAAGCCCTTGATGGATTTTCTTCAAGCGGGTTAGCGCCAC	4920	QY	5941	ACTATCATCATGATTTAGTTAATGCGATGATCTATATGATTTCTTTGAAACATAGATGTC	6000
Db	4861	TTCCCAATCAGCTAAATGTTTAAAGCCCTTGATGGATTTTCTTCAAGCGGGTTAGCGCCAC	4920	Db	5941	ACTATCATCATGATTTAGTTAATGCGATGATCTATATGATTTCTTTGAAACATAGATGTC	6000
QY	4921	GCAACAAGAAATGTGAAGCGGAGAGAAATGATCAGGATCGGGGTAGGGGATGG	4980	QY	6001	ATTTATCTGGATATAAAGATGGCGTTTAAACCTACTTTTGCAATTTTGTATATCTTTCT	6060
Db	4921	GCAACAAGAAATGTGAAGCGGAGAGAAATGATCAGGATCGGGGTAGGGGATGG	4980	Db	6001	ATTTATCTGGATATAAAGATGGCGTTTAAACCTACTTTTGCAATTTTGTATATCTTTCT	6060
QY	4981	AGTGAATACTTATCAAGAACTTTTGTGGTAATATCAACATAAATCTCAGGAGGAACGA	5040	QY	6061	TCTAATACATATGATCAATACACTTTTGTGTTTTTAAAGAAATTTAAAACTTATTTCAAAC	6120
Db	4981	AGTGAATACTTATCAAGAACTTTTGTGGTAATATCAACATAAATCTCAGGAGGAACGA	5040	Db	6061	TCTAATACATATGATCAATACACTTTTGTGTTTTTAAAGAAATTTAAAACTTATTTCAAAC	6120
QY	5041	GGAATACACATCATGCGGAGGTAAACAGTTCTTGGACCGGTTTCCACCTCCAACAACCTCAAC	5100	QY	6121	ATCGATCAGATTTTACTTTTGTTCCTATTTGCTACATTTATAGGCTCAGACTTTTGT	6180
Db	5041	GGAATACACATCATGCGGAGGTAAACAGTTCTTGGACCGGTTTCCACCTCCAACAACCTCAAC	5100	Db	6121	ATCGATCAGATTTTACTTTTGTTCCTATTTGCTACATTTATAGGCTCAGACTTTTGT	6180
QY	5101	AGGCCATCTCTAATCTAAGTACTCAGCACTAGCTATTTCTTGATGATTTCTTTTGTGGTT	5160	QY	6181	TTTCGAGATCAGATATCAGATCAATCACTTGTCTTTGAAAGACGTTGGAAACGCTCTTCT	6240
Db	5101	AGGCCATCTCTAATCTAAGTACTCAGCACTAGCTATTTCTTGATGATTTCTTTTGTGGTT	5160	Db	6181	TTTCGAGATCAGATATCAGATCAATCACTTGTCTTTGAAAGACGTTGGAAACGCTCTTCT	6240
QY	5161	GGGGTGTACATTTGGTCTGTCAAGTGTATTTGCTGAGGAGATCAAAACCAATGAGCT	5220	QY	6241	TTTCCACGATGTTTCTCGTGGGTGGGGTCCATCTTTTGGGACCACTGTCGGTAGAGGCAT	6300
Db	5161	GGGGTGTACATTTGGTCTGTCAAGTGTATTTGCTGAGGAGATCAAAACCAATGAGCT	5220	Db	6241	TTTCCACGATGTTTCTCGTGGGTGGGGTCCATCTTTTGGGACCACTGTCGGTAGAGGCAT	6300
QY	5221	ATATCCAAAGGCTAAATTTTGGAGCTCAAGGAAAGGTATGGTTTATAAAACTATCTTTT	5280	QY	6301	CTTGAACGATAGCTTTTCTTTTATCGCAATGATGGAATTTGTAAGGACCATCTTCTCTTT	6360
Db	5221	ATATCCAAAGGCTAAATTTTGGAGCTCAAGGAAAGGTATGGTTTATAAAACTATCTTTT	5280	Db	6301	CTTGAACGATAGCTTTTCTTTTATCGCAATGATGGAATTTGTAAGGACCATCTTCTCTTT	6360
				QY	6361	CTACTGCTCTTTCGATGAAAGTGCACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCG	6420

Db 6361 |||||CTACTGCTCTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCG 6420  
QY 6421 ATATTAACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTTGAGACTGTATCTTTGA 6480  
Db 6421 ATATTAACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTTGAGACTGTATCTTTGA 6480  
QY 6481 TATTTCTTGAGTAGACGAGAGTGTCGTCTCCACCATGTGGGATCTAGATATCACATC 6540  
Db 6481 TATTTCTTGAGTAGACGAGAGTGTCGTCTCCACCATGTGGGATCTAGATATCACATC 6540  
QY 6541 AATCCACTTGTCTTGAAGACGTGTTGGAACGTCTTCTTTTCCACGATGTTCTCGTG 6600  
Db 6541 AATCCACTTGTCTTGAAGACGTGTTGGAACGTCTTCTTTTCCACGATGTTCTCGTG 6600  
QY 6601 GTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCC 6660  
Db 6601 GTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCTTTCC 6660  
QY 6661 TATCGCAATGATGGCAATTTGTAGAACCAATCTTCTTCTTCTACTGTCTTGTGAGAA 6720  
Db 6661 TATCGCAATGATGGCAATTTGTAGAACCAATCTTCTTCTTCTACTGTCTTGTGAGAA 6720  
QY 6721 GACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCTTTGTGAAGA 6780  
Db 6721 GACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCTTTGTGAAGA 6780  
QY 6781 TCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGTATTTCTTGGAGTAGACGAG 6840  
Db 6781 TCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGTATTTCTTGGAGTAGACGAG 6840  
QY 6841 TGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTTCTTGTGAAGAG 6900  
Db 6841 TGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTTCTTGTGAAGAG 6900  
QY 6901 TGGTTGGAACGCTTCTTTTCCACGATGTTCTCGTGGTGGGGTCCATCTTTGGGAC 6960  
Db 6901 TGGTTGGAACGCTTCTTTTCCACGATGTTCTCGTGGTGGGGTCCATCTTTGGGAC 6960  
QY 6961 CACTGTCGTAGAGGCATCTTGAACGATAGCCCTTCTTTATCGCAATGATGGCATTTGT 7020  
Db 6961 CACTGTCGTAGAGGCATCTTGAACGATAGCCCTTCTTTATCGCAATGATGGCATTTGT 7020  
QY 7021 AGAAGCCATCTTCTTCTTCTACTGTCCTTTCGATGAAGTGACAGATAGCTGGGCAATGA 7080  
Db 7021 AGAAGCCATCTTCTTCTTCTACTGTCCTTTCGATGAAGTGACAGATAGCTGGGCAATGA 7080  
QY 7081 ATCCGAGGAGTTTCCGATATACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTT 7140  
Db 7081 ATCCGAGGAGTTTCCGATATACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTT 7140  
QY 7141 CTGAGACTGTATCTTTGTATTTCTTGGAGTAGACGAGAGTGTCGTCTCCACCATGTTGG 7200  
Db 7141 CTGAGACTGTATCTTTGTATTTCTTGGAGTAGACGAGAGTGTCGTCTCCACCATGTTGG 7200  
QY 7201 GGATCTAGATATCACATCAATCACTTGTGTTGAAGCTGTGGTGGAAAGTCTTCTTTT 7260  
Db 7201 GGATCTAGATATCACATCAATCACTTGTGTTGAAGCTGTGGTGGAAAGTCTTCTTTT 7260  
QY 7261 CCACGATGTTCTCGTGGTGGGGTCCATCTTGGGACCACTGTCGGTAGAGGCATCTT 7320  
Db 7261 CCACGATGTTCTCGTGGTGGGGTCCATCTTGGGACCACTGTCGGTAGAGGCATCTT 7320  
QY 7321 GAACGATAGCCTTCTTCTTATCGCAATGATGGCATTTGTAGAGCCATCTTCTTTCTA 7380  
Db 7321 GAACGATAGCCTTCTTCTTATCGCAATGATGGCATTTGTAGAGCCATCTTCTTTCTA 7380  
QY 7381 CTGCTCTTTGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATA 7440  
Db 7381 CTGCTCTTTGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATA 7440  
QY 7441 TTACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATAT 7500

Db 7441 TTACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATAT 7500  
QY 7501 TCTTGGAGTAGACGAGAGTGTCGTCTCCACCATGTGGGATCCACTAGTCTTAGAGCG 7560  
Db 7501 TCTTGGAGTAGACGAGAGTGTCGTCTCCACCATGTGGGATCCACTAGTCTTAGAGCG 7560  
QY 7561 GCCGCCACCGCGGTGGAGCT 7580  
Db 7561 GCCGCCACCGCGGTGGAGCT 7580  
RESULT 2  
ADL71846  
ID ADL71846 standard; DNA; 1746 BP.  
XX  
AC ADL71846;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Arabidopsis thaliana OBP3 antisense DNA #7.  
XX  
KW Transgenic plant; Dof transcription factor; ocs binding factor;  
plant size; plant stature; root growth; plant; gene; ds; OBF;  
OBF binding protein; OBP3; SOB1; mouse-ear cress.  
XX  
OS Arabidopsis thaliana.  
XX  
PN US2004045055-A1.  
XX  
PD 04-MAR-2004.  
XX  
PF 28-AUG-2003; 2003US-00650249.  
XX  
PR 28-AUG-2002; 2002US-0406657P.  
XX  
PI (UNIW ) UNIV WASHINGTON.  
XX  
Neff MM;  
XX  
WPI; 2004-225757/21.  
XX  
New transgenic plant cell, useful in producing plants with altered size  
and stature and with normal and healthy root growth.  
XX  
Claim 38; SEQ ID NO 18; 53pp; English.  
XX  
The invention relates to a transgenic plant transformed by a Dof  
transcription factor, OBF (ocs binding factor) binding protein (OBP3).  
OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic  
acid and polypeptides are useful in producing transgenic plants with  
altered size and stature and with normal and healthy root growth. The  
present sequence is Arabidopsis thaliana OBP3 antisense DNA.  
XX  
SQ Sequence 1746 BP; 560 A; 304 C; 304 G; 578 T; 0 U; 0 Other;  
Query Match 19.8%; Score 1503; DB 12; Length 1746;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1743; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 4436 AACAGAGAAGCAATCCAGATCCAGATCGAAATCTAGCGTGGTCTCGACTGATAATACTACT 4495  
Db 1 AACAGAGAAGCAATCCAGATCCAGATCGAAATCTAGCGTGGTCTCGACTGATAATACTACT 60  
QY 4496 AGTACTTTCATCACTTACTTCTCGCCCAAGTTACTCAAACTAGCAAGTTTCATAGCTAC 4555  
Db 61 AGTACTTTCATCACTTACTTCTCGCCCAAGTTACTCAAACTAGCAAGTTTCATAGCTAC 120  
QY 4556 GGTCAAAATCCCGAGAGTTAAATCCAACTTGCCCATCTTGCCCTCTCTCCAAAGCCTTGA 4615  
Db 121 GGTCAAAATCCCGAGAGTTAAATCCAACTTGCCCATCTTGCCCTCTCTCCAAAGCCTTGA 180  
QY 4616 GATTACAATTCAAGCAACACTCGATTAGATTTTGGTGGNACTCAATAGCAACATGATA 4675



Db 181 GATTACAANTCAAGCAACACTGGATTAGATTTTGGTGGAACTCAAAATGAAGCAACATGATA 240  
Qy 4676 AGTGGTATGAGTTCTAGTGGTGGGATCTTTGGATGATGGAGAATACTCCATCAACAACAA 4735  
Db 241 AGTGGTATGAGTTCTAGTGGTGGGATCTTTGGATGATGGAGAATACTCCATCAACAACAA 300  
Qy 4736 GCTCAGCAATCCCTTTCTTGATCAACACTACCGGATTTGGTGCATCTTCAACCGCGTTA 4795  
Db 301 GCTCAGCAATCCCTTTCTTGATCAACACTACCGGATTTGGTGCATCTTCAACCGCGTTA 360  
Qy 4796 TATCAATTTACTAGAGGTAAGGAGGTGTTAATCAAGGTGATTTCTCAACAGAGAAGTAGT 4855  
Db 361 TATCAATTTACTAGAGGTAAGGAGGTGTTAATCAAGGTGATTTCTCAACAGAGAAGTAGT 420  
Qy 4856 GATTATTTCCAATCAGCTAATGTTTAAAGCCCTTGATGGATTTTCTTCAGGCGGGTTAGC 4915  
Db 421 GATTATTTCCNATCAGCTAATGTTTAAAGCCCTTGATGGATTTTCTTCAGGCGGGTTAGC 480  
Qy 4916 GCCACGCAACCAAGAAATGTGAAGCGGAAGAGAATGATCAGGATCGGGGTAGGATGGG 4975  
Db 481 GCCACGCAACCAAGAAATGTGAAGCGGAAGAGAATGATCAGGATCGGGGTAGGATGGG 540  
Qy 4976 GATGAGTCAATTAATTATCAAGAACTTTTGGGTAAATATCAACATAAACTCAGGCAGG 5035  
Db 541 GATGAGTCAATTAATTATCAAGAACTTTTGGGTAAATATCAACATAAACTCAGGCAGG 600  
Qy 5036 AACGAGGAATACACATCATCGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCCAACAC 5095  
Db 601 AACGAGGAATACACATCATCGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCCAACAC 660  
Qy 5096 TCAACAGGCCATCTCTCAATTCCTAATTAAGTACTCAGCACTAGCTATTTCTTGATGATCTTT 5152  
Db 661 TCAACAGGCCATCTCTCAATTCCTAATTAAGTACTCAGCACTAGCTATTTCTTGATGATCTTT 720  
Qy 5153 TGTGTGGTGGGTGATACATTTGGTCTGTCATCGGAGTTATTGCTGAGGAAGATCAAAAC 5212  
Db 721 TGTGTGGTGGGTGATACATTTGGTCTGTCATCGGAGTTATTGCTGAGGAAGATCAAAAC 780  
Qy 5213 ATGCAGCTATATCCAAAGGCTAATTTTGAGGCTCAAAAGAAAGGTATGTTTATAAACA 5272  
Db 781 ATGCAGCTATATCCAAAGGCTAATTTTGAGGCTCAAAAGAAAGGTATGTTTATAAACA 840  
Qy 5273 TCTTTTGTATCTTTTAAAAAGATCTTCAAAAGTGTAGTATGTTTATGTTGGCTTCGT 5332  
Db 841 TCTTTTGTATCTTTTAAAAAGATCTTCAAAAGTGTAGTATGTTTATGTTGGCTTCGT 900  
Qy 5333 GATATTTATGTTTATAGAAATTTGCTTATATATTGCTATATAGAGGTGTGGG 5392  
Db 901 GATATTTATGTTTATAGAAATTTGCTTATATATTGCTATATAGAGGTGTGGG 960  
Qy 5393 ATATGTATGAATTCAGAGGTGTATGTTGGAACCTTTTGTGTGTTTCATTGAATAATCAT 5452  
Db 961 ATATGTATGAATTCAGAGGTGTATGTTGGAACCTTTTGTGTGTTTCATTGAATAATCAT 1020  
Qy 5453 CGAATTTCAATTTCTTGAGAGCCCATATATGAGACATTTGAGACATCTATAGACATATAT 5512  
Db 1021 CGAATTTCAATTTCTTGAGAGCCCATATATGAGACATTTGAGACATCTATAGACATATAT 1080  
Qy 5513 GTAATGTATATTAACCGTACTTAAGTCGAATTTTATGACCACAAGTAATAAATATATGCCG 5572  
Db 1081 GTAATGTATATTAACCGTACTTAAGTCGAATTTTATGACCACAAGTAATAAATATATGCCG 1140  
Qy 5573 AATGTATCATGCTAATATPCAGATTTTAAACTATTTTCCAAATATAAACAATTTTCTCTT 5632  
Db 1141 AATGTATCATGCTAATATPCAGATTTTAAACTATTTTCCAAATATAAACAATTTTCTCTT 1200  
Qy 5633 TCGTCCAACTATATATCTTATCTGATCTTATTTCTTTTCTTTTAAATTCCTTTTCC 5692  
Db 1201 TCGTCCAACTATATATCTTATCTGATCTTATTTCTTTTCTTTTAAATTCCTTTTCC 1260  
Qy 5693 TTTCCCAAGACACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5752  
Db 1261 TTTCCCAAGACACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1320

Qy 5753 TAAACCACGAGAAATATGCACCTAAATTCAGACTAAATCCCCAAATTTTCAGAAATTTATG 5812  
Db 1321 TAAACCACGAGAAATATGCACCTAAATTCAGACTAAATCCCCAAATTTTCAGAAATTTATG 1380  
Qy 5813 TATTTTTCGGAATTTAAATATTTGTTCACAAATCAATAATGGCCAACTAACTAATTTGAAAAGA 5872  
Db 1381 TATTTTTCGGAATTTAAATATTTGTTCACAAATCAATAATGGCCAACTAACTAATTTGAAAAGA 1440  
Qy 5873 CAATGGAATGATGAAACCATGCATAATCTCTCAAGTCTCAACCTATGAAGAAATCATGTA 5932  
Db 1441 CAATGGAATGATGAAACCATGCATAATCTCTCAAGTCTCAACCTATGAAGAAATCATGTA 1500  
Qy 5933 ACCAATAGACTATCATCATGATTAAGTAAATGCATGATCTATAATCTTTTGTGAACATA 5992  
Db 1501 ACCAATAGACTATCATCATGATTAAGTAAATGCATGATCTATAATGATATCTTTGTGAACATA 1560  
Qy 5993 GATATGTCATTTATCTCGATATAAAGATGGGGTTTTTAACCTACTTTGCAATTTTGTGTAT 6052  
Db 1561 GATATGTCATTTATCTGGATATAAGATGGGGTTTTTAACCTACTTTTGTGTAT 1620  
Qy 6053 ATCTTTCTCTTAATACATATGATCAATACACTTTTGTGTTTTTAAAGAAATTTAAACCTTA 6112  
Db 1621 ATCTTTCTCTTAATACATATGATCAATACACTTTTGTGTTTTTAAAGAAATTTAAACCTTA 1680  
Qy 6113 TTTCAACATCGATCAGATTTTACTTTTGTTCATATTCAGTACATTTATAGGCTCAC 6172  
Db 1681 TTTCAACATCGATCAGATTTTACTTTTGTTCATATTCAGTACATTTATAGGCTCAC 1740  
Qy 6173 ACTTTT 6178  
Db 1741 ACTTTT 1746

## RESULT 3

ABQ73047/c

ID ABQ73047 standard; DNA; 10078 BP.

XX ABQ73047;

XX 24-SEP-2002 (first entry)

XX Tomato anthocyanin 1 (ANT1) related plasmid pAG3202 SEQ ID NO:3.

XX Tomato; ANT1; anthocyanin 1; Lycopersicon esculentum; plant; colour;

XX gene; ds.

XX Lycopersicon esculentum.

XX Synthetic.

XX WO200255658-A2.

XX 18-JUL-2002.

XX 29-OCT-2001; 2001WO-US050638.

XX 30-OCT-2000; 2000US-0244685P.

XX (EXEL-) EXELIXIS PLANT SCI INC.

XX Connors K, Mathews HV, Liu A;

XX WPI; 2002-557819/59.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

New isolated polynucleotide derived from tomato, useful for producing an Anthocyanin 1 phenotype in plants, particularly for modifying e.g. leaf color, flower color or fruit color in plants.

Example 1; Page 41-46; 53pp; English.

The present invention describes tomato anthocyanin 1 (ANT1). The ANT1 polynucleotide can be used for modifying the expression of a native plant gene, particularly for producing an anthocyanin 1 phenotype in plants,

CC which is responsible for many red and blue colours in plants. The  
CC polynucleotide is useful for modifying e.g. leaf colour, flower colour or  
CC fruit colour in plants. The present sequence represents the plasmid  
CC pAG3202 which is used in an example from the present invention for the  
CC generation of plants with an AN1 phenotype by transformation with an  
CC activation tagging construct  
XX  
SQ Sequence 10078 BP; 2533 A; 2568 C; 2625 G; 2333 T; 0 U; 19 Other;

Query Match 18.5%; Score 1402; DB 6; Length 10078;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6179 GTTTCGGATCTAGATATACATCAATCCACTTCTTTGAAGACGTTGGTGGAAAGCTCTTC 6238  
DB 4249 GTTTCGGATCTAGATATACATCAATCCACTTCTTTGAAGACGTTGGTGGAAAGCTCTTC 4190  
QY 6239 TTTTTCACGATGTTCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGC 6298  
DB 4189 TTTTTCACGATGTTCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGC 4130  
QY 6299 ATCTTGAAGCATAGCCCTTCTTTATCGCAATGCGCATTTGTAGAAGCCATCTTCTCTT 6358  
DB 4129 ATCTTGAAGCATAGCCCTTCTTTATCGCAATGCGCATTTGTAGAAGCCATCTTCTCTT 4070  
QY 6359 TTCTACTGTCCTTTCCATGAAGTGAAGATAGCTGGGCAATGGAATCCGAGGAGTTTCC 6418  
DB 4069 TTCTACTGTCCTTTCCATGAAGTGAAGATAGCTGGGCAATGGAATCCGAGGAGTTTCC 4010  
QY 6419 CGATATTACCCCTTTGTTGAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTT 6478  
DB 4009 CGATATTACCCCTTTGTTGAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTT 3950  
QY 6479 GATATTCTTGGAGTAGACAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCACA 6538  
DB 3949 GATATTCTTGGAGTAGACAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCACA 3890  
QY 6539 TCAATCCAATGCTTTGAAGAAGTGTGGTGGAAAGCTTCTTTTCCACGATGTTCTCGT 6598  
DB 3889 TCAATCCAATGCTTTGAAGAAGTGTGGTGGAAAGCTTCTTTTCCACGATGTTCTCGT 3830  
QY 6599 GGGTGGGGTCCATCTTTGGGACCACTGTCGTTAGAGGATCTTGAACGATAGCCCTTCC 6658  
DB 3829 GGGTGGGGTCCATCTTTGGGACCACTGTCGTTAGAGGATCTTGAACGATAGCCCTTCC 3770  
QY 6659 TTTATCGCAATGATGCAATTTGTAGAAGCATCTTCTTTCTACTGTCCTTTCCGATGAA 6718  
DB 3769 TTTATCGCAATGATGCAATTTGTAGAAGCATCTTCTTTCTACTGTCCTTTCCGATGAA 3710  
QY 6719 GTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATTAACCTTTGTTGAAA 6778  
DB 3709 GTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATTAACCTTTGTTGAAA 3650  
QY 6779 AGTCTCAATAGCCCTCTGCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAG 6838  
DB 3649 AGTCTCAATAGCCCTCTGCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAG 3590  
QY 6839 AGTGTGCTGCTCCACCATGTTGGGATCTAGATATCACAATCAATCCACTTGTCTTGAAGA 6898  
DB 3589 AGTGTGCTGCTCCACCATGTTGGGATCTAGATATCACAATCAATCCACTTGTCTTGAAGA 3530  
QY 6899 CGTGGTTGGAAGCTCTTCTTTTCCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGG 6958  
DB 3529 CGTGGTTGGAAGCTCTTCTTTTCCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGG 3470  
QY 6959 ACCACTGTCGGTAGAGGCATCTTGAAGCATAGCTTCTTTATCGCAATGATGCGATTT 7018  
DB 3469 ACCACTGTCGGTAGAGGCATCTTGAAGCATAGCTTCTTTATCGCAATGATGCGATTT 3410  
QY 7019 GTAGAAGCCATCTTCTTTCTACTGTCCTTTTCGATGAAGTGCACAGATAGCTGGCAATG 7078  
DB 3409 GTAGAAGCCATCTTCTTTCTACTGTCCTTTTCGATGAAGTGCACAGATAGCTGGCAATG 3350

QY 7079 GAATCCGAGGAGTTTCCGATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTC 7138  
DB 3349 GAATCCGAGGAGTTTCCGATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTC 3290  
QY 7139 TTCTCAGACTGTATCTTTGATATTTCTTGAGTAGACGAGAGTGTCTGCTCCACCATGTT 7198  
DB 3289 TTCTCAGACTGTATCTTTGATATTTCTTGAGTAGACGAGAGTGTCTGCTCCACCATGTT 3230  
QY 7199 GGGGATCTAGATATCACAATCAATCCACTTCTTTGAAGACGTTGGTGGAAAGCTCTTCTTT 7258  
DB 3229 GGGGATCTAGATATCACAATCAATCCACTTCTTTGAAGACGTTGGTGGAAAGCTCTTCTTT 3170  
QY 7259 TTCCACGATGTTCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGCATC 7318  
DB 3169 TTCCACGATGTTCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGCATC 3110  
QY 7319 TTGAACGATAGCTTCTTCTTTATCGCAATGATGCGATTTGTAGAAGCCATCTTCTTTTTC 7378  
DB 3109 TTGAACGATAGCTTCTTCTTTATCGCAATGATGCGATTTGTAGAAGCCATCTTCTTTTTC 3050  
QY 7379 TACTGTCTTCTCGATGAAGTGAAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGA 7438  
DB 3049 TACTGTCTTCTCGATGAAGTGAAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGA 2990  
QY 7439 TATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGAT 7498  
DB 2989 TATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGAT 2930  
QY 7499 ATTCTTGGAGTAGACAGAGTGTCTGTCCTCCACCATGTTGGGATCCACTAGTCTTAGAG 7558  
DB 2929 ATTCTTGGAGTAGACAGAGTGTCTGTCCTCCACCATGTTGGGATCCACTAGTCTTAGAG 2870  
QY 7559 CGGCCGCCACCGCGTGGAGCT 7580  
DB 2869 CGGCCGCCACCGCGTGGAGCT 2848  
  
RESULT 4  
ADV39061/c  
ID ADV39061 standard; DNA; 17511 BP.  
XX  
AC ADV39061;  
XX  
XX 24-FEB-2005 (first entry)  
XX  
DE Plant gene activation vector-related DNA sequence pHR-AT - SEQ ID 13.  
XX  
KW gene activation vector; genetic analysis; ds.  
XX  
OS Unidentified.  
XX  
XX WO2004106520-A1.  
XX  
PD 09-DEC-2004.  
XX  
XX 28-MAY-2004; 2004WO-JP007789.  
XX  
XX 30-MAY-2003; 2003JP-00153985.  
XX  
XX (RIKE ) RIKEN KK.  
XX  
XX Muranaka T, Seki H;  
XX  
XX WPI; 2005-021291/02.  
XX  
PT Novel plant gene active vector comprising hairy-root induction gene, and  
PT enhancer and/or promoter that functions in plant, useful for activating  
PT plant gene forming hairy root.  
XX  
PS Claim 4; SEQ ID NO 13; 54pp; Japanese.  
XX  
XX The invention comprises a plant gene activation vector that contains a  
CC hairy-root induction gene and enhancer and/or promoter which can function

CC in a plant. The vector of the invention is useful for analyzing a plant  
CC gene. The present DNA sequence is claimed in the invention.  
XX  
SQ Sequence 17511 BP; 4453 A; 4201 C; 4511 G; 4344 T; 0 U; 2 Other;  
Query Match 18.0%; Score 1367; DB 14; Length 17511;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6179 GTTTCGGATCTAGATATACATCAATCCACTTGTCTTGAAGACGTGGTTGGAACGCTCTTC 6238  
DB 17503 GTTTCGGATCTAGATATACATCAATCCACTTGTCTTGAAGACGTGGTTGGAACGCTCTTC 17444  
QY 6239 TTTTCCACGATGTCCTCGTGGGTGGGGTGCATCTTTGGGACCACTGTCGGTAGAGC 6298  
DB 17443 TTTTCCACGATGTCCTCGTGGGTGGGGTGCATCTTTGGGACCACTGTCGGTAGAGC 17384  
QY 6299 ATCTTGAAGCATAGCCTTTCCCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCTT 6358  
DB 17383 ATCTTGAAGCATAGCCTTTCCCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCTT 17324  
QY 6359 TTCTACTGTCCTTTCCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCC 6418  
DB 17323 TTCTACTGTCCTTTCCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCC 17264  
QY 6419 CGATATTACCTTTGTGTAAGAGTCTCAATAGCCCTCGGTCTTCTGAGACTGTATCTTT 6478  
DB 17263 CGATATTACCTTTGTGTAAGAGTCTCAATAGCCCTCGGTCTTCTGAGACTGTATCTTT 17204  
QY 6479 GATATTCTGGAGTAGACAGAGTGTGCTGCTCCACCATGTTGGGATCTAGATATCACA 6538  
DB 17203 GATATTCTGGAGTAGACAGAGTGTGCTGCTCCACCATGTTGGGATCTAGATATCACA 17144  
QY 6539 TCAATCCACTTGTGTAAGACGTGGTTGGAAAGTCTTCTTTTCCACGATGTCCTCGT 6598  
DB 17143 TCAATCCACTTGTGTAAGACGTGGTTGGAAAGTCTTCTTTTCCACGATGTCCTCGT 17084  
QY 6599 GGGTGGGGTCCATCTTTTGGGACCACTGTCGGTAGAGGCACTTGAACGATGACCTTTCC 6658  
DB 17083 GGGTGGGGTCCATCTTTTGGGACCACTGTCGGTAGAGGCACTTGAACGATGACCTTTCC 17024  
QY 6659 TTTATCGCAATGATGCAATTTGTAGAAGCCATCTTCTTCTACTGTCCTTTCCGATGAA 6718  
DB 17023 TTTATCGCAATGATGCAATTTGTAGAAGCCATCTTCTTCTACTGTCCTTTCCGATGAA 16964  
QY 6719 GTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCTTTGTTGAAA 6778  
DB 16963 GTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCTTTGTTGAAA 16904  
QY 6779 AGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTTCTGGAGTAGACGAG 6838  
DB 16903 AGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTTCTGGAGTAGACGAG 16844  
QY 6839 AGTCTGCTCCACCATGTTGGGATCTAGATATCACATCAATCCACTTGTCTTGAAGA 6898  
DB 16843 AGTCTGCTCCACCATGTTGGGATCTAGATATCACATCAATCCACTTGTCTTGAAGA 16784  
QY 6899 CGTGGTTGGAACGCTCTTTTCCACGATGTTTCCCTGCGGTGGGGTCCATCTTTGGG 6958  
DB 16783 CGTGGTTGGAACGCTCTTTTCCACGATGTTTCCCTGCGGTGGGGTCCATCTTTGGG 16724  
QY 6959 ACCACTGTCGGTAGAGGCAATCTTGAACGATAGCCTTTCTTTATCCGAATGATGGCATTT 7018  
DB 16723 ACCACTGTCGGTAGAGGCAATCTTGAACGATAGCCTTTCTTTATCCGAATGATGGCATTT 16664  
QY 7019 GTAGAAGCATCTTCTCTTCTACTCTCTTCTGTAAGAGTGACATAGCTGGGCAATG 7078  
DB 16663 GTAGAAGCATCTTCTCTTCTACTCTCTTCTGTAAGAGTGACATAGCTGGGCAATG 16604  
QY 7079 GAATCCGAGGAGGTTTCCCGATATTACCTTTCTTGAAGAGTCTCAATAGCCCTCTGGTC 7138  
DB 16603 GAATCCGAGGAGGTTTCCCGATATTACCTTTCTTGAAGAGTCTCAATAGCCCTCTGGTC 16544

QY 7139 TTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTT 7198  
DB 16543 TTCTGAGACTGTATCTTTGATATTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTT 16484  
QY 7199 GGGGATCTAGATATCACATCAATCCACTTGTCTTGAAGACGTGGTTGGAACGCTCTTCTTT 7258  
DB 16483 GGGGATCTAGATATCACATCAATCCACTTGTCTTGAAGACGTGGTTGGAACGCTCTTCTTT 16424  
QY 7259 TTCCACGATGTCCTCGTGGGTGGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCAATC 7318  
DB 16423 TTCCACGATGTCCTCGTGGGTGGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCAATC 16364  
QY 7319 TTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 7378  
DB 16363 TTGAACGATAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 16304  
QY 7379 TACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGA 7438  
DB 16303 TACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGA 16244  
QY 7439 TATTACCTTTGTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGAT 7498  
DB 16243 TATTACCTTTGTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGAT 16184  
QY 7499 ATTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCC 7545  
DB 16183 ATTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTGGGATCC 16137

RESULT 5  
ADV39062/c  
ID ADV39062 standard; DNA; 18987 BP.  
XX AC ADV39062;  
XX DT 24-FEB-2005 (first entry)  
XX DE Plant gene activation vector-related DNA sequence pHR-AT-GFP - SEQ ID 14.  
XX KW gene activation vector; genetic analysis; ds.  
XX OS Unidentified.  
XX PN WO2004106520-A1.  
XX PD 09-DEC-2004.  
XX PF 28-MAY-2004; 2004WO-JP007789.  
XX PR 30-MAY-2003; 2003JP-00153985.  
XX PA (RIKE ) RIKEN KK.  
XX PI Muranaka T, Seki H;  
XX DR WPI; 2005-021291/02.  
XX PT Novel plant gene active vector comprising hairy-root induction gene, and  
PT enhancer and/or promoter that functions in plant, useful for activating  
PT plant gene forming hairy root.  
XX PS Claim 5; SEQ ID NO 14; 54pp; Japanese.  
XX CC The invention comprises a plant gene activation vector that contains a  
CC hairy-root induction gene and enhancer and/or promoter which can function  
CC in a plant. The vector of the invention is useful for analyzing a plant  
CC gene. The present DNA sequence is claimed in the invention.  
XX SQ Sequence 18987 BP; 4871 A; 4609 C; 4863 G; 4642 T; 0 U; 2 Other;

Query Match 18.0%; Score 1367; DB 14; Length 18987;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	6179	GTTCGGATCTAGATATCAATCAATCCACTTCTGCTTTGAAGACGTGGTGGAACTGCTTC	6238
Db	18979	GTTCGGATCTAGATATCAATCAATCCACTTCTGCTTTGAAGACGTGGTGGAACTGCTTC	18920
QY	6239	TTTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGC	6298
Db	18919	TTTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGC	18860
QY	6299	ATCTTGAAGATAGCTTTCTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCTT	6358
Db	18859	ATCTTGAAGATAGCTTTCTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCTT	18800
QY	6359	TTCTACTGCTCTTTCCATGAAGTGAAGATGATGCTGGGCAATGGAATCCGAGGAGGTTCC	6418
Db	18799	TTCTACTGCTCTTTCCATGAAGTGAAGATGATGCTGGGCAATGGAATCCGAGGAGGTTCC	18740
QY	6419	CGATATTACCTTTGTTGAAGTCTCAATAGCCCTCTCGTCTTCTGAGACTGTATCTTT	6478
Db	18739	CGATATTACCTTTGTTGAAGTCTCAATAGCCCTCTCGTCTTCTGAGACTGTATCTTT	18680
QY	6479	GATATTCTTGGAGTACAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCACA	6538
Db	18679	GATATTCTTGGAGTACAGAGTGTCTGCTCCACCATGTTGGGATCTAGATATCACA	18620
QY	6539	TCAATCCACTTGCTTTGAAGACGTGGTGGAAAGTCTCTTTTCCACGATGTCCTGT	6598
Db	18619	TCAATCCACTTGCTTTGAAGACGTGGTGGAAAGTCTCTTTTCCACGATGTCCTGT	18560
QY	6599	GGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGATCTTGAACGATAGCTTTCC	6658
Db	18559	GGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGATCTTGAACGATAGCTTTCC	18500
QY	6659	TTTATCGCAATGATGCAATTTGTAGAAGCATCTCTCTTTTCTACTGCTCTTTTGAAG	6718
Db	18499	TTTATCGCAATGATGCAATTTGTAGAAGCATCTCTCTTTTCTACTGCTCTTTTGAAG	18440
QY	6719	GTGACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCGATATPACCTTTGTTGAAA	6778
Db	18439	GTGACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCGATATPACCTTTGTTGAAA	18380
QY	6779	AGTCTCAATAGCCCTGCTCTCTGAGACTGTATCTTTGATATCTTGGAGTAGACGAG	6838
Db	18379	AGTCTCAATAGCCCTGCTCTCTGAGACTGTATCTTTGATATCTTGGAGTAGACGAG	18320
QY	6839	AGTGTGCTGCTCCACCATGTTGGGATCTAGATATCACATCAATCCACTTGTCTTGAAGA	6898
Db	18319	AGTGTGCTGCTCCACCATGTTGGGATCTAGATATCACATCAATCCACTTGTCTTGAAGA	18260
QY	6899	CGTGGTTGGAACGTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGG	6958
Db	18259	CGTGGTTGGAACGTCTCTTTTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGG	18200
QY	6959	ACCAGTCTGGTAGAGGATCTTGAACGATAGCTTTCCCTTTATCGCAATGATGGCATTT	7018
Db	18199	ACCAGTCTGGTAGAGGATCTTGAACGATAGCTTTCCCTTTATCGCAATGATGGCATTT	18140
QY	7019	GTAGAAGCCATCTTCTCTTCTACTGCTCTCTTTCGATGAAGTGCACAGATAGCTGGCAATG	7078
Db	18139	GTAGAAGCCATCTTCTCTTCTACTGCTCTCTTTCGATGAAGTGCACAGATAGCTGGCAATG	18080
QY	7079	GAATCCGAGGAGTTTCCGATATPACCTTTTGTGAAAAGTCTCAATAGCCCTCTGCTTC	7138
Db	18079	GAATCCGAGGAGTTTCCGATATPACCTTTTGTGAAAAGTCTCAATAGCCCTCTGCTTC	18020
QY	7139	TTCTGAGACTGTATCTTTGATATCTTTGAGTAGAGGATGTCGTCTCCACCATGTT	7198
Db	18019	TTCTGAGACTGTATCTTTGATATCTTTGAGTAGAGGATGTCGTCTCCACCATGTT	17960
QY	7199	GGGGATCTAGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGAACTGCTCTTT	7258
Db	17959	GGGGATCTAGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGAACTGCTCTTT	17900
QY	7259	TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGATC	7318
Db	17899	TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGATC	17840
QY	7319	TTGAAAGATAGCTTTCTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCTTTTC	7378
Db	17839	TTGAAAGATAGCTTTCTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCTTTTC	17780
QY	7379	TACTGTCTCTTTTCGATGAAGTGCACAGATAGCTGGCAATGGAATCCGAGGAGGTTTCCCGA	7438
Db	17779	TACTGTCTCTTTTCGATGAAGTGCACAGATAGCTGGCAATGGAATCCGAGGAGGTTTCCCGA	17720
QY	7439	TATTACCTTTTGTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGAT	7498
Db	17719	TATTACCTTTTGTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGAT	17660
QY	7499	ATTCTTGGAGTAGACAGAGTGTCTGCTCCACCATGTTGGGATCC	7545
Db	17659	ATTCTTGGAGTAGACAGAGTGTCTGCTCCACCATGTTGGGATCC	17613
RESULT 6			
AAA88401/c			
ID	AAA88401	standard; DNA; 1361 BP.	
XX	AAA88401;		
AC	AAA88401;		
DT	09-JAN-2001	(first entry)	
XX	4X	CaMV 35S enhancer construct.	
DE	CaMV; 35S; enhancer; trait-associated gene identification;		
KW	activation tagging; fruit; transgenic plant; ss.		
XX	Cauliflower mosaic virus.		
XX	Key	Location/Qualifiers	
FT	repeat_region	1..1354	
FT	/tag= a	/repeat_type= TANDEM	
FT	/note= "4 CaMV 35S enhancer units"		
FT	repeat_unit	1..339	
FT	/tag= b	/note= "CaMV 35S enhancer unit 1"	
FT	enhancer	1..129	
FT	/tag= c	/note= "CaMV 35S enhancer AluI-EcoRV fragment"	
FT	enhancer	130..331	
FT	/tag= d	/note= "129 bp fragment of the CaMV sequence"	
FT	enhancer	333..339	
FT	/tag= e	/note= "CaMV 35S enhancer unit 2"	
FT	repeat_unit	340..678	
FT	/tag= f	/note= "additional 7 bp not associated with 35S enhancer"	
FT	enhancer	340..541	
FT	/tag= g	/note= "CaMV 35S enhancer AluI-EcoRV fragment"	
FT	enhancer	542..670	
FT	/tag= h	/note= "129 bp fragment of the CaMV sequence"	
FT	enhancer	672..678	
FT	/tag= i	/note= "additional 7 bp not associated with 35S enhancer"	
FT	repeat_unit	679..1017	
FT	/tag= j	/note= "CaMV 35S enhancer unit 3"	
FT	enhancer	679..880	
FT	/tag= k	/note= "CaMV 35S enhancer AluI-EcoRV fragment"	
FT	enhancer	881..1009	
FT	/tag= l		

FT enhancer /note= "129 bp fragment of the CaMV sequence"  
FT 1011. .1017  
FT /tag= m  
FT /note= "Additional 7 bp not associated with 35S enhancer"  
FT repeat\_unit 1018. .1354  
FT /tag= a  
FT /note= "CaMV 35S enhancer unit 4"  
FT enhancer 1018. .1219  
FT /tag= n  
FT /note= "CaMV 35S enhancer AluI-EcoRV fragment"  
FT enhancer 1220. .1348  
FT /tag= o  
FT /note= "129 bp fragment of the CaMV sequence"  
FT enhancer 1350. .1354  
FT /tag= p  
FT /note= "Additional 6 bp not associated with 35S enhancer"  
XX WO200053794-A2.  
XX 14-SEP-2000.  
XX 09-MAR-2000; 2000WO-US006298.  
XX 12-MAR-1999; 99US-0124232P.  
XX (AGRI-) AGRITOPE INC.  
XX Wagner R, Mathews H, Liu XL, Waggoner WJ;  
XX WPI; 2000-594336/56.  
XX  
XX Identifying genes associated with a desired trait for isolating and  
XX characterizing the genes comprises using an enhancer element which  
XX enhances gene expression and stably integrates into the plant genome.  
XX  
XX Claim 4; Fig 5; 42pp; English.  
XX  
XX The present sequence is that of a 4X cauliflower mosaic virus (CaMV) 35S  
XX enhancer sequence preferred for use in the method of the invention. It  
XX includes 4 repeats of 202 bp AluI-EcoRV fragments of the 35S enhancer,  
XX 129 bp of the CaMV sequence associated with each tandem Alu-EcoRV repeat,  
XX and an additional 7 bp repeated sequence, which does not appear in the  
XX 35S enhancer region of the native CaMV genome. This 4X CaMV 35S enhancer  
XX element can be used in a method for identifying genes associated with a  
XX desired trait in a fruit-bearing plant. The method involves: transforming  
XX plant cells with an activation tagging vector comprising an element which  
XX functions to enhance gene expression and has the ability to integrate  
XX into the plant genome in a manner effective to enhance expression of  
XX native plant genes, selecting transformed plant cells, regenerating  
XX transformed plant cells to yield mature plants, selecting plants having a  
XX desired trait, identifying, isolating and characterizing genes the  
XX transcription of which has been enhanced, and confirming the contribution  
XX of the modified expression of each identified gene to the desired trait.  
XX The desired trait may be increased resistance to fungal, bacterial or  
XX viral pathogens, insects, modifications in flower size, flower number,  
XX flower pigmentation and shape, modified leaf number, leaf pigmentation  
XX and shape, modified seed number, pattern or distribution of leaves and  
XX flowers, modified stem length between nodes, root mass or root  
XX development characteristics or increased drought, salt and antibiotic  
XX tolerance. Plants having short life cycles are transformed, as  
XX exemplified by dwarf varieties of tomato  
XX  
XX Sequence 1361 BP; 447 A; 332 C; 297 G; 285 T; 0 U; 0 Other;  
SQ  
Query Match 17.9%; Score 1360; DB 3; Length 1361;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6179 GTTTCGGATCTAGATATCACATCAATCCACTTCTTTGAAGACGTGGTGGAAACGTCCTC 6238  
DB 1360 GTTTCGGATCTAGATATCACATCAATCCACTTCTTTGAAGACGTGGTGGAAACGTCCTC 1301  
QY 6239 TTTTTCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGGACCACCTGTCGGTAGAGGC 6298

Db 1300 TTTTTCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGGACCACCTGTCGGTAGAGGC 1241  
QY 6299 ATCTTGAAGGATAGCTCTTCTTTATCCAAATGATGGCATTTGTAGAAGCCATCTTCCCTT 6358  
Db 1240 ATCTTGAAGGATAGCTCTTCTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCCCTT 1181  
QY 6359 TTCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCC 6418  
Db 1180 TTCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCC 1121  
QY 6419 CGATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTT 6478  
Db 1120 CGATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTT 1061  
QY 6479 GATATTCTTGGAGTAGACGAGAGTGTCTGCCACCATGTTGGGGATCTAGATATCACA 6538  
Db 1060 GATATTCTTGGAGTAGACGAGAGTGTCTGCCACCATGTTGGGGATCTAGATATCACA 1001  
QY 6539 TCAATCCACTTCTTGAAGACGTGTTGGAACGTCTTCTTTTCCACGATGTTCTCGT 6598  
Db 1000 TCAATCCACTTCTTGAAGACGTGTTGGAACGTCTTCTTTTCCACGATGTTCTCGT 941  
QY 6599 GGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAAACGATAGCCCTTCC 6658  
Db 940 GGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAAACGATAGCCCTTCC 881  
QY 6659 TTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTTCTACTGTCTTTTCGATGAA 6718  
Db 880 TTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTTCTACTGTCTTTTCGATGAA 821  
QY 6719 GTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTTACCTTTGTTGAAA 6778  
Db 820 GTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTTACCTTTGTTGAAA 761  
QY 6779 AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAG 6838  
Db 760 AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAG 701  
QY 6839 AGTGTGTGCTCCACCATGTTGGGATCTAGATATCAGATCAATCCACTTGTCTTTGAAAGA 6898  
Db 700 AGTGTGTGCTCCACCATGTTGGGATCTAGATATCAGATCAATCCACTTGTCTTTGAAAGA 641  
QY 6899 CGTGGTGGAAAGCTCTTCTTTTCCACGATGTTCTCTGCGGTGGGGTCCATCTTTGGG 6958  
Db 640 CGTGGTGGAAAGCTCTTCTTTTCCACGATGTTCTCTGCGGTGGGGTCCATCTTTGGG 581  
QY 6959 ACCACTGTCTGAGAGGCATCTTGAACGATAGCCCTTCTTTATTCGCAATGATGGCATTT 7018  
Db 580 ACCACTGTCTGAGAGGCATCTTGAACGATAGCCCTTCTTTATTCGCAATGATGGCATTT 521  
QY 7019 GTAGAAGCCATCTTCTTTTCTACTGTCTTTCGATGAAGTAGACATAGCTGGGCAATG 7078  
Db 520 GTAGAAGCCATCTTCTTTTCTACTGTCTTTCGATGAAGTAGACATAGCTGGGCAATG 461  
QY 7079 GAATCCGAGAGGTTTCCCGATATTTACCTTTGTAAGTAGCTCAATAGCCCTCTGGTC 7138  
Db 460 GAATCCGAGAGGTTTCCCGATATTTACCTTTGTAAGTAGCTCAATAGCCCTCTGGTC 401  
QY 7139 TTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTT 7198  
Db 400 TTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTT 341  
QY 7199 GGGGATCTAGATATCACATCAATCCACTTGTCTTTGAAGAGCGTGGTGGAAACGTCCTTTT 7258  
Db 340 GGGGATCTAGATATCACATCAATCCACTTGTCTTTGAAGAGCGTGGTGGAAACGTCCTTTT 281  
QY 7259 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGTGAGAGGCATC 7318  
Db 280 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGTGAGAGGCATC 221  
QY 7319 TTGAACGATAGCCCTTTCTTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTTTC 7378

Db 220 TTGAACGATAGCCTTTCCTTTATCGCAATGATGGCAATTTGTAGAGCCATCTCCTTTTC 161

Qy 7379 TACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGA 7438

Db 160 TACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGA 101

Qy 7439 TATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGAT 7498

Db 100 TATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGAT 41

Qy 7499 ATTCTTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTG 7538

Db 40 ATCTTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTTG 1

RESULT 7

ADL71845

ID ADL71845 standard; DNA; 1235 BP.

XX

AC ADL71845;

XX

DT 20-MAY-2004 (first entry)

XX

DE Arabidopsis thaliana OBP3 antisense DNA #6.

XX

KW Transgenic plant; Dof transcription factor; ocs binding factor;

KW plant size; plant stature; root growth; plant; gene; db; OBP;

KW OBF binding protein; OBP3; SOB1; mouse-ear cress.

XX

OS Arabidopsis thaliana.

XX

PN US2004045055-A1.

XX

PD 04-MAR-2004.

XX

PF 28-AUG-2003; 2003US-00650249.

XX

PR 28-AUG-2002; 2002US-0406657P.

XX

PA (UNIW ) UNIV WASHINGTON.

XX

PI Neff MM;

XX

DR WPI; 2004-225757/21.

XX

PT New transgenic plant cell, useful in producing plants with altered size

PT and stature and with normal and healthy root growth.

XX

PS Claim 37; SEQ ID NO 17; 53pp; English.

XX

CC The invention relates to a transgenic plant transformed by a Dof

CC transcripction factor, OBF (ocs binding factor) binding protein (OBP3).

CC OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic

CC acid and polypeptides are useful in producing transgenic plants with

CC altered size and stature and with normal and healthy root growth. The

CC present sequence is Arabidopsis thaliana OBP3 antisense DNA.

XX

SQ Sequence 1235 BP; 408 A; 222 C; 179 G; 426 T; 0 U; 0 Other;

Query Match 16.3%; Score 1235; DB 12; Length 1235;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3045 CAACCAAGAACGATGACGTATATGATTGACTTGCAAAAATAAGCAAAACAAATACCTGTT 3104

Db 1 CAACCAAGAACGATGACGTATATGATTGACTTGCAAAAATAAGCAAAACAAATACCTGTT 60

Qy 3105 CAAATCGACACTTAATTCCAAAAAGGTTAGTAATAGTAAGAAGGCTTTTATTATGAAA 3164

Db 61 CAAATCGACACTTAATTCCAAAAAGGTTAGTAATAGTAAGAAGGCTTTTATTATGAAA 120

Qy 3165 ACAAAGAANAATAAGAGCCTTAAGCAATGATGAAAATTTGAAAGAGAAAAAGAGCATTTG 3224

Db 121 ACAAAGAANAATAAGAGCCTTAAGCAATGATGAAAATTTGAAAGAGAAAAAGAGCATTTG 180

Qy 3225 TTATAGAAAAGAAAAAGAGAGAGTAAAGAGAAATTAAGAAACACAAATAAAATTAACAA 3284

Db 181 TTATAGAAAAGAAAAAGAGAGAGTAAAGAGAAATTAAGAAACACAAATAAAATTAACAA 240

Qy 3285 AGGAAACTTCATTTCTCTCTTTATCCCAATTCAGCTCCCTCTCTCTCTCTCTCTCTCTCT 3344

Db 241 AGGAAACTTCATTTCTCTCTTTATCCCAATTCAGCTCCCTCTCTCTCTCTCTCTCTCTCT 300

Qy 3345 CTCTCTCTCTAGATCAATTTCTTTCTTATGATGATGATTAATCCACCAATATCTCGGACC 3404

Db 301 CTCTCTCTCTAGATCAATTTCTTTCTTATGATGATGATTAATCCACCAATATCTCGGACC 360

Qy 3405 TCTTACCTAAAAAGGATACAAGTAAGAGATTCAAAGATGGTTTTCTCATCTCTCTCGAGT 3464

Db 361 TCTTACCTAAAAAGGATACAAGTAAGAGATTCAAAGATGGTTTTCTCATCTCTCTCGAGT 420

Qy 3465 AATCAGTTCGATTCCTCCAAATTTGGCAGCAGGTAAATAATCAGTTTATGATATTTGCTAGAT 3524

Db 421 AATCAGTTCGATTCCTCCAAATTTGGCAGCAGGTAAATAATCAGTTTATGATATTTGCTAGAT 480

Qy 3525 GTTCTCTGATTCGTTCTCTTTCTCTCAAGCTCGATCAAGATTTATGAAAATTTGATGAGA 3584

Db 481 GTTCTCTGATTCGTTCTCTTTCTCTCAAGCTCGATCAAGATTTATGAAAATTTGATGAGA 540

Qy 3585 TTTTGTTCGACAAAATTCCTAGCTATTGTGGACGGCATATATATPATTAATTAATTAATTC 3644

Db 541 TTTTGTTCGACAAAATTCCTAGCTATTGTGGACGGCATATATATPATTAATTAATTAATTC 600

Qy 3645 TTAGTTGATTAACCCCTTTTCTTCTCTCTCGAATATACGAAATATATAAGAT 3704

Db 601 TTAGTTGATTAACCCCTTTTCTTCTCTCTCGAATATACGAAATATATAAGAT 660

Qy 3705 GATTTCAATTTTGGTCTTTTCTTCTCAAGCTTTTAAATAATTTCTTCTAGTTGA 3764

Db 661 GATTTCAATTTTGGTCTTTTCTTCTCAAGCTTTTAAATAATTTCTTCTAGTTGA 720

Qy 3765 TAAAAACCTTTTCTTGTCTCTCAAGGGCTTATGATATAATGTTTTTCTTACAGGAT 3824

Db 721 TAAAAACCTTTTCTTGTCTCTCAAGGGCTTATGATATAATGTTTTTCTTACAGGAT 780

Qy 3825 AATTTTCTCTTTGGTTAGATTTTTCACCGCCATGGAATATCACTTCAAAAATAAAAA 3884

Db 781 AATTTTCTCTTTGGTTAGATTTTTCACCGCCATGGAATATCACTTCAAAAATAAAAA 840

Qy 3885 GTTTTAAAGTTACTATGACTTTTAAATCTGAGTTATTTATCCATTTTCTTTTTCGAGCTTTGT 3944

Db 841 GTTTTAAAGTTACTATGACTTTTAAATCTGAGTTATTTATCCATTTTCTTTTTCGAGCTTTGT 900

Qy 3945 TGA AAAA ACTATATAATTAATCTGCAATTTCTGTCAAAGTAGTCACAAATTTTATCTATTTTC 4004

Db 901 TGA AAAA ACTATATAATTAATTAATCTGCAATTTCTGTCAAAGTAGTCACAAATTTTATCTATTTTC 960

Qy 4005 TTTTGTCTCCGACCAATGTTTCAAACCTCGAATCCCTTTCGTTTAAAGTTGTTTCTGCTTTA 4064

Db 961 TTTTGTCTCCGACCAATGTTTCAAACCTCGAATCCCTTTCGTTTAAAGTTGTTTCTGCTTTA 1020

Qy 4065 TTATAAACCTGAAACTTAATTAGTACAAATTTATGTTTAAATGATGAGCAAGGGAACCAACATC 4124

Db 1021 TTATAAACCTGAAACTTAATTAGTACAAATTTATGTTTAAATGATGAGCAAGGGAACCAACATC 1080

Qy 4125 AGCTAGAATGTGTCACACTGACCAAGAACCTTAATAATTTACTTCAGGAGCTCTCATCAC 4184

Db 1081 AGCTAGAATGTGTCACACTGACCAAGAACCTTAATAATTTACTTCAGGAGCTCTCATCAC 1140

Qy 4185 CACCGACTTCTCAGGTTGAGGTTGAGTCAAGCTAGAGTGAATTCATTTGGTGGACGTTG 4244

Db 1141 CACCGACTTCTCAGGTTGAGGTTGAGTCAAGCTAGAGTGAATTCATTTGGTGGACGTTG 1200

Qy 4245 CTCGGATCGCAAAAGTCCCATTTGCCCTGAAGCAGCT 4279

Db 1201 CTCGGATCGCAAAAGTCCCATTTGCCCTGAAGCAGCT 1235



RESULT 8	
ADL71848	
ID	ADL71848 standard; DNA; 1062 BP.
XX	
AC	ADL71848;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Arabidopsis thaliana OBP3 antisense DNA #9.
XX	
KW	Transgenic plant; Dof transcription factor; ocs binding factor;
KW	plant size; plant stature; root growth; plant; gene; db; OBF;
KW	OBF binding protein; OBP3; SOB1; mouse-ear cress.
XX	
OS	Arabidopsis thaliana.
XX	
PN	US2004045055-A1.
XX	
PD	04-MAR-2004.
XX	
PF	28-AUG-2003; 2003US-00650249.
XX	
PR	28-AUG-2002; 2002US-0406657P.
XX	
PA	(UNIW ) UNIV WASHINGTON.
XX	
PI	Neff MM;
XX	
DR	WPI; 2004-225757/21.
XX	
PT	New transgenic plant cell, useful in producing plants with altered size
PT	and stature and with normal and healthy root growth.
XX	
PS	Claim 40; SEQ ID NO 20; 53pp; English.
XX	
CC	The invention relates to a transgenic plant transformed by a Dof
CC	transcription factor, OBF (ocs binding factor) binding protein (OBP3).
CC	OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic
CC	acid and polypeptides are useful in producing transgenic plants with
CC	altered size and stature and with normal and healthy root growth. The
CC	present sequence is Arabidopsis thaliana OBP3 antisense DNA.
XX	
SQ	Sequence 1062 BP; 348 A; 159 C; 155 G; 400 T; 0 U; 0 Other;
Query Match	14.0%; Score 1062; DB 12; Length 1062;
Best Local Similarity	100.0%; Pred. No. 7e-308;
Matches 1062; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	5117 TAAGTACTCAGCACTAGCTATTCTTGATGATTCCTTTTGGTGGGTGTACATTGGTG 5176
DB	1 TAGGTACTCAGCACTAGCTATTCTTGATGATTCCTTTTGGTGGGTGTACATTGGTG 60
QY	5177 CTTGTGATCGGAGTTATGCTGAGGAAGATCAACCATCGAGCTATATCCAAAGGCTAAT 5236
DB	61 CTTGTGATCGGAGTTATGCTGAGGAAGATCAACCATCGAGCTATATCCAAAGGCTAAT 120
QY	5237 TTTGAGGCTCAAGGAAGATGCTGTTATAAACTATCTTTTGGTCTTTTAAAGATCT 5296
DB	121 TTTGAGGCTCAAGGAAGATGCTGTTATAAACTATCTTTTGGTCTTTTAAAGATCT 180
QY	5297 TCAAGGTGTCAGTATGTTTATTTGGTGGCTTCCTGGTGATATTTATGTTTATAGAAATTT 5356
DB	181 TCAAGGTGTCAGTATGTTTATTTGGTGGCTTCCTGGTGATATTTATGTTTATAGAAATTT 240
QY	5357 GGTCTTATATATTTGGCTATATATAGAGGTGGTGATGATGATGAATTCAGAGTTGAT 5416
DB	241 GGTCTTATATATTTGGCTATATATAGAGGTGGTGATGATGATGAATTCAGAGTTGAT 300
QY	5417 GTTGGAAACCTTTTGTGCTTCAATTAATCAATCGAATTCCTCAATTTCTTTGGAGACC 5476
DB	301 GTTGGAAACCTTTTGTGCTTCAATTAATCAATCGAATTCCTCAATTTCTTTGGAGACC 360

QY	5477 CATTATGAGACATTGAGACATCTATAGAACATATATGTAATGTAATTAACGTACTTAA 5536
DB	361 CATTATGAGACATTGAGACATCTATAGAACATATATGTAATGTAATTAACGTACTTAA 420
QY	5537 GTCGAATTTTATGACCAAGTAAATTAATGTCGGAATGTACATGCTAAATATCGAGTTT 5596
DB	421 GTCGAATTTTATGACCAAGTAAATTAATGTCGGAATGTACATGCTAAATATCGAGTTT 480
QY	5597 AAACATATTTTCCAAATATAACAATATTTCTCTTTGTCGCAACTATATATCTCTATT 5656
DB	481 AAACATATTTTCCAAATATAACAATATTTCTCTTTGTCGCAACTATATATCTCTATT 540
QY	5657 CTGATCTTATTTCTCTTTTAAATTCCTTTTCCCAAGACACAAAAA 5716
DB	541 CTGATCTTATTTCTCTTTTAAATTCCTTTTCCCAAGACACAAAAA 600
QY	5717 AATACAGAAACGAAAAAGAGATTTTAAATTCATAACCCACGAGAAATATGACCTTA 5776
DB	601 AATACAGAAACGAAAAAGAGATTTTAAATTCATAACCCACGAGAAATATGACCTTA 660
QY	5777 AATCAGACTAATCCCCCAATTTTCAGAAATTTATGATTTTTCGATTTAATATGTGT 5836
DB	661 AATTCAGACTAATCCCCCAATTTTCAGAAATTTATGATTTTTCGATTTAATATGTGT 720
QY	5837 TCACAAATCATATGCGCAACTAACTAATTTGAAAGACAAATGGAATGAAACCATGCA 5896
DB	721 TCACAAATCATATGCGCAACTAATTTGAAAGACAAATGGAATGAAACCATGCA 780
QY	5897 TAATCTCTCAAGTCTCAACCTATGAGAAATCATGTAAACCAATAGACTATCATCATGATTA 5956
DB	781 TAATCTCTCAAGTCTCAACCTATGAGAAATCATGTAAACCAATAGACTATCATCATGATTA 840
QY	5957 GTTAATGATGATCTATATATGATTTCTTTGAAACATAGATATGCTATTATCTGGATATA 6016
DB	841 GTTAATGATGATCTATATATGATTTCTTTGAAACATAGATATGCTATTATCTGGATATA 900
QY	6017 AGATGGCGTTTAACTCTTGGCAATTTTCTGTTATATCTTTCTTAATACATATGATC 6076
DB	901 AGATGGCGTTTAACTCTTGGCAATTTTCTGTTATATCTTTCTTAATACATATGATC 960
QY	6077 AATACACTTTTGTGTTTAAAGAAATTTAAAACTTTATTTCAAACATCGATCATTTTTTA 6136
DB	961 AATACACTTTTGTGTTTAAAGAAATTTAAAACTTTATTTCAAACATCGATCATTTTTTA 1020
QY	6137 CTTTGTGTTTCCATATGACTACATTTATAGGCTCACACTTTT 6178
DB	1021 CTTTGTGTTTCCATATGACTACATTTATAGGCTCACACTTTT 1062
RESULT 9	
ADL71847	
ID	ADL71847 standard; DNA; 1058 BP.
XX	
AC	ADL71847;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Arabidopsis thaliana OBP3 antisense DNA #8.
XX	
KW	Transgenic plant; Dof transcription factor; ocs binding factor;
KW	plant size; plant stature; root growth; plant; gene; db; OBF;
KW	OBF binding protein; OBP3; SOB1; mouse-ear cress.
XX	
OS	Arabidopsis thaliana.
XX	
PN	US2004045055-A1.
XX	
PD	04-MAR-2004.
XX	
PF	28-AUG-2003; 2003US-00650249.
XX	
PR	28-AUG-2002; 2002US-0406657P.



```
XX (UNIW ) UNIV WASHINGTON.
XX Neff MM;
XX WPI; 2004-225757/21.
XX
XX New transgenic plant cell, useful in producing plants with altered size
XX and stature and with normal and healthy root growth.
XX
XX Claim 39; SEQ ID NO 19; 53pp; English.
XX
XX The invention relates to a transgenic plant transformed by a Dof
XX transcription factor, OBP (ocs binding factor) binding protein (OBP3).
XX OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic
XX acid and polypeptides are useful in producing transgenic plants with
XX altered size and stature and with normal and healthy root growth. The
XX present sequence is Arabidopsis thaliana OBP3 antisense DNA.
XX
XX Sequence 1058 BP; 355 A; 176 C; 139 G; 388 T; 0 U; 0 Other;
XX
XX Query Match 14.0%; Score 1058; DB 12; Length 1058;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-306;
XX Matches 1058; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 3045 CAACCAAGAACGATGACGTATATGATGACTTGCAGAAATTAAGCAAAATACCTGTT 3104
DB 1 CAACCAAGAACGATGACGTATATGATGACTTGCAGAAATTAAGCAAAATACCTGTT 60
QY 3105 CAATCGACACATTAATCCAAAAGGTTAGTAAAGTAAAGAGGCTTTTATATGAAA 3164
DB 61 CAATCGACACATTAATCCAAAAGGTTAGTAAAGTAAAGAGGCTTTTATATGAAA 120
QY 3165 ACAGAAAGAAATAAGAGGCTTAAGAGATGATCAAAATTTGAAAGAGAAAAGAGCATG 3224
DB 121 ACAGAAAGAAATAAGAGGCTTAAGAGATGATCAAAATTTGAAAGAGAAAAGAGCATG 180
QY 3225 TTATGAAAAAGAAAAAGAGAGAGTAAAGAGATTAAGAAACACAAATAAATAACAA 3284
DB 181 TTATGAAAAAGAAAAAGAGAGAGTAAAGAGATTAAGAAACACAAATAAATAACAA 240
QY 3285 AGGAACTTCATTTCTCTTTATCCCATTCAGCTCCCTCTCTCTCTCTCTCTCTCTCT 3344
DB 241 AGGAACTTCATTTCTCTTTATCCCATTCAGCTCCCTCTCTCTCTCTCTCTCTCTCT 300
QY 3345 CTCTCTCTCTAGATCAATTTCTTCTATGATGATGATATCAACCATATCTCGACC 3404
DB 301 CTCTCTCTCTAGATCAATTTCTTCTATGATGATGATATCAACCATATCTCGACC 360
QY 3405 TCTTACCTAAAAGGATACAAGTAAAGATTCAGAGATGTTTCTCATCTCTTCCAGTG 3464
DB 361 TCTTACCTAAAAGGATACAAGTAAAGATTCAGAGATGTTTCTCATCTCTTCCAGTG 420
QY 3465 AATCAGTTCGATCCGAAATTTGGCAGCAGGTAAATACAGTTATGATATTGCTAGAT 3524
DB 421 AATCAGTTCGATCCGAAATTTGGCAGCAGGTAAATACAGTTATGATATTGCTAGAT 480
QY 3525 GTTTCGTGATTCGTTCTCTTTCTCCAGCTCGATCAAGATTTATGAAAATTTGATGAGA 3584
DB 481 GTTTCGTGATTCGTTCTCTTTCTCCAGCTCGATCAAGATTTATGAAAATTTGATGAGA 540
QY 3585 TTTTGTTCGACAAAATTCCTAGCTATTTGGAGCGCATATATATTTACTTATGAATATTC 3644
DB 541 TTTTGTTCGACAAAATTCCTAGCTATTTGGAGCGCATATATATTTACTTATGAATATTC 600
QY 3645 TTAGTTGATTAACCCCTTTTCTCTCTCTCTCGATATACGAAATATATAAGAT 3704
DB 601 TTAGTTGATTAACCCCTTTTCTCTCTCTCTCGATATACGAAATATATAAGAT 660
QY 3705 GATTTCAAATTTGGTCTTTTCTTCTACTCTCAAGACTTTTAAAAAATTTATCTTAGTTCA 3764
DB 661 GATTTCAAATTTGGTCTTTTCTTCTACTCTCAAGACTTTTAAAAAATTTATCTTAGTTCA 720
```

```
QY 3765 TAAAAACCTTTTCTTCTCTCCAGGGCTTATGATATAATGTTTCTTACAGATT 3824
DB 721 TAAAAACCTTTTCTTCTCTCCAGGGCTTATGATATAATGTTTCTTACAGATT 780
QY 3825 AATTTTCTCTTTGGTTAGATTTTACACCGCCATGGAATATACACTTCAAAAATAAAAA 3884
DB 781 AATTTTCTCTTTGGTTAGATTTTACACCGCCATGGAATATACACTTCAAAAATAAAAA 840
QY 3885 GTTTAAAGTTACTATGACCTTAAATCTGAGTTATTTATCAATTTCTTTTGCAGCTTTGT 3944
DB 841 GTTTAAAGTTACTATGACCTTAAATCTGAGTTATTTATCAATTTCTTTTGCAGCTTTGT 900
QY 3945 TCAAAAACCTATAATTAATCTGCAATTTCTGTCAAGTAGTCAAAATTTTATCTATTTTC 4004
DB 901 TCAAAAACCTATAATTAATTAATCTGCAATTTCTGTCAAGTAGTCAAAATTTTATCTATTTTC 960
QY 4005 TTTTGTCTCCGACCAATGTTTCAAACTCGAACTCTTCTTCTTAAAGTTCTTTCTGCTTTA 4064
DB 961 TTTTGTCTCCGACCAATGTTTCAAACTCGAACTCTTCTTCTTAAAGTTCTTTCTGCTTTA 1020
QY 4065 TTATAAACCTGAACTAAATTAATGATCAAAATTTATGTTAAT 4102
DB 1021 TTATAAACCTGAACTAAATTAATGATCAAAATTTATGTTAAT 1058
XX
XX RESULT 10
XX ADL71840
XX ID ADL71840 standard; DNA; 888 BP.
XX AC ADL71840;
XX XX
XX DT 20-MAY-2004 (first entry)
XX XX
XX DE Arabidopsis thaliana OBP3 antisense DNA #1.
XX XX
XX KW Transgenic plant; Dof transcription factor; ocs binding factor;
XX KW plant size; plant stature; root growth; plant; gene; ds; OBF;
XX KW OBF binding protein; OBP3; SOB1; mouse-ear cress.
XX XX
XX OS Arabidopsis thaliana.
XX XX
XX PN US2004045055-A1.
XX XX
XX PD 04-MAR-2004.
XX XX
XX PF 28-AUG-2003; 2003US-00650249.
XX XX
XX PR 28-AUG-2002; 2002US-0406657P.
XX XX
XX PA (UNIW ) UNIV WASHINGTON.
XX XX
XX PI Neff MM;
XX XX
XX WPI; 2004-225757/21.
XX XX
XX PT New transgenic plant cell, useful in producing plants with altered size
XX PT and stature and with normal and healthy root growth.
XX XX
XX PS Claim 26; SEQ ID NO 12; 53pp; English.
XX XX
XX CC The invention relates to a transgenic plant transformed by a Dof
XX CC transcription factor, OBP (ocs binding factor) binding protein (OBP3).
XX CC OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic
XX CC acid and polypeptides are useful in producing transgenic plants with
XX CC altered size and stature and with normal and healthy root growth. The
XX CC present sequence is Arabidopsis thaliana OBP3 antisense DNA.
XX XX
XX SQ Sequence 888 BP; 263 A; 193 C; 197 G; 235 T; 0 U; 0 Other;
XX
XX Query Match 11.7%; Score 888; DB 12; Length 888;
XX Best Local Similarity 100.0%; Pred. No. 6.8e-256;
XX Matches 888; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 4232 ATGTTGGAACGTCCTCGGATCGCAAAAGTCCCATTCGCTGAGCAGCTCTAAATTTGCCCT 4291  
Db 1 ATGTTGGAACGTCCTCGGATCGCAAAAGTCCCATTCGCTGAGCAGCTCTAAATTTGCCCT 60  
QY 4292 AGATGTGACTCAACCAATACTAAGTTCTGTTACTTTCAATAACTATAGCCTTACTCAACT 4351  
Db 61 AGATGTGACTCAACCAATACTAAGTTCTGTTACTTTCAATAACTATAGCCTTACTCAACT 120  
QY 4352 CCCCATTCTTGCAAAACATGTCGTGCTATTTGGAACGCTGGCGGTTCTTGAAGAAATGTT 4411  
Db 121 CCCCATTCTTGCAAAACATGTCGTGCTATTTGGAACGCTGGCGGTTCTTGAAGAAATGTT 180  
QY 4412 CTTGTTGAGGAGGCTTTAGGAGGAACAAGAGAAGCAATCCAGATCGAAATCTACGGTC 4471  
Db 181 CTTGTTGAGGAGGCTTTAGGAGGAACAAGAGAAGCAATCCAGATCGAAATCTACGGTC 240  
QY 4472 GTGGTCTCGACTGATAAATACTACTAGTACTTCTCATCTACTTCTTCGCCCCAAGTTACTCA 4531  
Db 241 GTGGTCTCGACTGATAAATACTACTAGTACTTCTCATCTACTTCTTCGCCCCAAGTTACTCA 300  
QY 4532 AACCTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATTCGAATCTGCCCATC 4591  
Db 301 AACCTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATTCGAATCTGCCCATC 360  
QY 4592 TTGCTCTCTCCAAAGCCTTGGAGTTACAAATTCAGCAACACACACACACACACACACAC 4651  
Db 361 TTGCTCTCTCCAAAGCCTTGGAGTTACAAATTCAGCAACACACACACACACACACACAC 420  
QY 4652 GGAATCTCAATAAGCAACATGATAGTGTATGATTTCTAGTGTGGGATCTTTGGATGCA 4711  
Db 421 GGAATCTCAATAAGCAACATGATAGTGTATGATTTCTAGTGTGGGATCTTTGGATGCA 480  
QY 4712 TGGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGA 4771  
Db 481 TGGAGAAATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGA 540  
QY 4772 TTGTTGCAATCTTCAACCGTTATATCCATTACTAGAAGTTAAGGAGGTCTTAATCAA 4831  
Db 541 TTGTTGCAATCTTCAACCGTTATATCCATTACTAGAAGTTAAGGAGGTCTTAATCAA 600  
QY 4832 GTGATTTCTCAACAGAAAGTAGTGTATTTCCAAATCAGCTAATGTTTAAAGCCTTGATG 4891  
Db 601 GTGATTTCTCAACAGAAAGTAGTGTATTTCCAAATCAGCTAATGTTTAAAGCCTTGATG 660  
QY 4892 GATTTTTCTTCAGCGGGGTTAGCGCCACGCAAAACAGAAATGTGAAGCGGGAAGAGAAAT 4951  
Db 661 GATTTTTCTTCAGCGGGGTTAGCGCCACGCAAAACAGAAATGTGAAGCGGGAAGAGAAAT 720  
QY 4952 GATCAGGATCGGGTAGGGATGCGGATGGAGTGAATTAATCTTATCAGAAACTTTTGGGT 5011  
Db 721 GATCAGGATCGGGTAGGGATGCGGATGGAGTGAATTAATCTTATCAGAAACTTTTGGGT 780  
QY 5012 AATATCAACATAAATCTCAGCAGGAACGAGGAATACACATCATGCGGAGGTAAACAGTTCT 5071  
Db 781 AATATCAACATAAATCTCAGCAGGAACGAGGAATACACATCATGCGGAGGTAAACAGTTCT 840  
QY 5072 TGGACCGGTTTACCTCCAACTCAACACTCAACAGGCCATCTCTCATTTCTAA 5119  
Db 841 TGGACCGGTTTACCTCCAACTCAACACTCAACAGGCCATCTCTCATTTCTAA 888

## RESULT 11

ADL71843

XX ID ADL71843 standard; DNA; 684 BP.

XX AC ADL71843;

XX ADL71843;

XX 20-MAY-2004 (first entry)

XX Arabidopsis thaliana OBP3 antisense DNA #4.

XX Transgenic plant; Dof transcription factor; ocs binding factor;  
XX plant size; plant stature; root growth; plant; gene; db; OBF;  
KW

KW OBF binding protein; OBP3; SOB1; mouse-ear cress.

XX Arabidopsis thaliana.

OS US2004045055-A1.

XX 04-MAR-2004.

XX 28-AUG-2003; 2003US-00650249.

XX 28-AUG-2002; 2002US-0406657P.

XX (UNIW ) UNIV WASHINGTON.

XX Neff WM;

XX WPI; 2004-225757/21.

XX New transgenic plant cell, useful in producing plants with altered size  
XX and stature and with normal and healthy root growth.

XX Claim 29; SEQ ID NO 15; 53pp; English.

XX The invention relates to a transgenic plant transformed by a Dof

XX transcription factor, OBF (ocs binding factor) binding protein (OBP3).

XX OBP3 is also known as SOB1. The transgenic plant cell and OBP3 nucleic

XX acid and polypeptides are useful in producing transgenic plants with

XX altered size and stature and with normal and healthy root growth. The

XX present sequence is Arabidopsis thaliana OBP3 antisense DNA.

XX Sequence 684 BP; 212 A; 145 C; 149 G; 178 T; 0 U; 0 Other;

SQ Query Match 9.0%; Score 684; DB 12; Length 684;

Best Local Similarity 100.0%; Pred. No. 6.2e-195; Indels 0;

Matches 684; Conservative 0; Mismatches 0;

QY 4436 AACAGAGAGCAAAATCCAGATCGAAATCTACGGTCTGCTCGACTGATAATACTACT 4495

Db 1 AACAGAGAGCAAAATCCAGATCGAAATCTACGGTCTGCTCGACTGATAATACTACT 60

QY 4496 AGTACTTCAATCACTTCTTCGCCCAAGTTACTCAAACTTAGCAAGTTTCATAGCTAC 4555

Db 61 AGTACTTCAATCACTTCTTCGCCCAAGTTACTCAAACTTAGCAAGTTTCATAGCTAC 120

QY 4556 GGTCAAAATCCCGAGGTTTAAATTCCAACTTGGTGGAACTCAAAATAGCAACATGATA 4615

Db 121 GGTCAAAATCCCGAGGTTTAAATTCCAACTTGGTGGAACTCAAAATAGCAACATGATA 180

QY 4616 GATTACAATTCAGCAACACTGGATTTAGATTTTGGTGGAACTCAAAATAGCAACATGATA 4675

Db 181 GATTACAATTCAGCAACACTGGATTTAGATTTTGGTGGAACTCAAAATAGCAACATGATA 240

QY 4676 AGTGGTATGAGTTCTAGTGGTGGGATCTTGGATGATGAGAAATACCTCCTCAACAACAA 4735

Db 241 AGTGGTATGAGTTCTAGTGGTGGGATCTTGGATGATGAGAAATACCTCCTCAACAACAA 300

QY 4736 GCTCAGCAATTCCTTCTTGTATCAACACTACCGGATTTGGTGGAACTCAAAATAGCAACATGATA 4795

Db 301 GCTCAGCAATTCCTTCTTGTATCAACACTACCGGATTTGGTGGAACTCAAAATAGCAACATGATA 360

QY 4796 TATCAATTAAGAGGTAAAGGAGGTGTTAATCAAGGTGATTTCTCAACAGAGAGTAGT 4855

Db 361 TATCAATTAAGAGGTAAAGGAGGTGTTAATCAAGGTGATTTCTCAACAGAGAGTAGT 420

QY 4856 GATTATTCGAATCAGCTAATGTTTAAAGCCTTGTAGGATTTTCTTTCAGCGGGGTAGG 4915

Db 421 GATTATTCGAATCAGCTAATGTTTAAAGCCTTGTAGGATTTTCTTTCAGCGGGGTAGG 480

QY 4916 GCCACGCAAAACAGAAATGTGAAGCGGGAAGAGAGATGATCAGATCGGGGTAGGATGGG 4975

Db 481 GCCACGCAAAACAGAAATGTGAAGCGGGAAGAGAGATGATCAGATCGGGGTAGGATGGG 540

QY 4976 GATGGAGTGAATAAATTTATCAAGAACTTTTGGGTAAATATCAACATAAACTCAGGAGG 5035



PI Broun PE;  
XX WPI; 2003-248221/24.  
DR P-PSDB; ADD30378.  
XX  
PT New plant transcription factor polynucleotides and polypeptides, useful  
in producing transgenic plants with commercially valuable properties,  
PT such as an alteration in a plant growth characteristic, e.g. growth rate  
or apomixis.  
XX  
PS Disclosure; SEQ ID NO 406; 454pp; English.  
XX  
XX The invention relates to a number of isolated Arabidopsis thaliana cDNA  
sequences and their encoded proteins which are especially transcription  
factor related cDNA's and proteins. The isolated or recombinant plant  
transcription factor polynucleotides and polypeptides are useful in  
CC producing transgenic plants with commercially valuable properties, i.e.  
CC modified or altered desirable traits as compared to a reference plant,  
CC such as an alteration in a plant growth characteristic, e.g. growth rate,  
CC germination rate of seeds, vigor of plants and seedlings, or leaf and  
CC flower senescence. Sequence information related to the polynucleotides  
CC and polypeptides can also be used in bioinformatic search methods. The  
CC transgenic plant is useful for growing a progeny plant from a parent  
CC plant. This sequence represents one of the cDNAs of the invention.  
XX  
SQ Sequence 795 BP; 236 A; 180 C; 175 G; 204 T; 0 U; 0 Other;  
  
Query Match 7.7%; Score 581; DB 10; Length 795;  
Best Local Similarity 100.0%; Pred. No. 3.5e-164;  
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 4232 ATGGTGGAAAGCTCTCGGATCGCAAAAGTCCCATTCCTGGAAGCAGCTCTAAATTGCCCT 4291  
Db 1 ATGGTGGAAAGCTCTCGGATCGCAAAAGTCCCATTCCTGGAAGCAGCTCTAAATTGCCCT 60  
  
QY 4292 AGATGTGACTCAACCAATCTAAGTTCGTCTTCAATAACTATAGACCTTACTCAACT 4351  
Db 61 AGATGTGACTCAACCAATCTAAGTTCGTCTTCAATAACTATAGACCTTACTCAACT 120  
  
QY 4352 CGCCATTTCTGCAAAACATGTCGTCTATTGACACGCTGGCGGTTCTTGGAGGAATGTT 4411  
Db 121 CGCCATTTCTGCAAAACATGTCGTCTATTGACACGCTGGCGGTTCTTGGAGGAATGTT 180  
  
QY 4412 CTGTTGGAGGAGCTTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 4471  
Db 181 CTGTTGGAGGAGCTTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATCTACGGTC 240  
  
QY 4472 GTGGTCTGACATGATTAATCTACTAGTCTTCTATCTACTTCTCGCCCAAGTTACTCA 4531  
Db 241 GTGGTCTGACATGATTAATCTACTAGTCTTCTATCTACTTCTCGCCCAAGTTACTCA 300  
  
QY 4532 AACCTTAGCAAGTTTTCATAGCTACGGTCAAAATCCCGAGGTTTAAATTCCACTTGGCCATC 4591  
Db 301 AACCTTAGCAAGTTTTCATAGCTACGGTCAAAATCCCGAGGTTTAAATTCCACTTGGCCATC 360  
  
QY 4592 TTGCTCTCTCTCAAGCCCTTGAGATTCAAAATTCAGCAACAACCTGGATTAGATTGGT 4651  
Db 361 TTGCTCTCTCTCAAGCCCTTGAGATTCAAAATTCAGCAACAACCTGGATTAGATTGGT 420  
  
QY 4652 GGAATCAATAGCAACATGATAGTGTATGAGTTCTAGTGGTGGATCTTGGATGCA 4711  
Db 421 GGAATCAATAGCAACATGATAGTGTATGAGTTCTAGTGGTGGATCTTGGATGCA 480  
  
QY 4712 TGAGAAATACCTCCATCACAAAGCTCAGCAATTCCTTTCTTGTATCAACACTACCGGA 4771  
Db 481 TGAGAAATACCTCCATCACAAAGCTCAGCAATTCCTTTCTTGTATCAACACTACCGGA 540  
  
QY 4772 TTGGTGAATCTTCAACCGGTTATATCCATTTACTAGAAG 4812  
Db 541 TTGGTGAATCTTCAACCGGTTATATCCATTTACTAGAAG 581

RESULT 14

ADI44304  
ID ADI44304 standard; DNA; 795 BP.  
XX  
AC ADI44304;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Plant transcription factor related polynucleotide #1741.  
XX  
KW transgenic; plant; enhanced tolerance to abiotic stress;  
KW glycosphosphate tolerance; hormone sensitivity; disease resistance;  
KW sugar sensing; flowering; flower structure; stem bifurcation;  
KW branching pattern; apical dominance; trichome; stem morphology;  
KW root growth; root hair; seed development; cell proliferation;  
KW cell differentiation; premature senescence; necrosis; plant size;  
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;  
KW plant anthocyanin; light response; shade avoidance; bioinformatic;  
KW transcription factor; ds.  
XX  
OS Unidentified.  
XX  
PN US2004019927-A1.  
XX  
PD 29-JAN-2004.  
XX  
PF 25-FEB-2003; 2003US-00374780.  
XX  
PR 18-APR-2001; 2001US-00837944.  
XX  
PA (SHER/) SHERMAN B K.  
PA (RIEC/) RIECHMANN J L.  
PA (JIAN/) JIANG C.  
PA (HEAR/) HEARD J E.  
PA (HAAR/) HAANG V.  
PA (CREE/) CREELMAN R A.  
PA (RATC/) RATCLIFFE O.  
PA (ADAM/) ADAM L J.  
PA (REUB/) REUBER T L.  
PA (KEDD/) KEDDIE J.  
PA (BROU/) BROUN P E.  
PA (DUBE/) DUBELL M L.  
PA (PINE/) PINEDA O.  
PA (YUGG/) YU G.  
XX  
PI Sherman BK, Riechmann JB, Jiang C, Heard JE, Haake V;  
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;  
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;  
XX  
DR WPI: 2004-132245/13.  
XX P-PSDB; ADI44305.  
XX  
PT New transgenic plant comprising a recombinant polynucleotide of any one  
PT of more than 500 nucleotide sequences, useful in bioinformatic search  
PT methods.  
XX  
PS Disclosure; SEQ ID NO 2767; 435pp; English.  
XX  
XX The invention describes a transgenic plant comprising a recombinant  
polynucleotide of any one of more than 500 nucleotide sequences fully  
CC defined in the specification or its complement. The method of the  
CC invention can be used to produce a plant having altered traits such as:  
CC enhanced tolerance to abiotic stress; glycosphosphate tolerance; hormone  
CC sensitivity; disease resistance; sugar sensing; early or late flowering;  
CC altered flower structure, change in stem bifurcations, altered branching  
CC pattern, reduced apical dominance, reduced trichome density; lack of  
CC trichomes; reduced ectopic trichome development; altered trichome  
CC development; increase in trichome number; altered stem morphology;  
CC increased root growth, increased root hairs; altered seed development;  
CC altered cell proliferation or cell differentiation; rapid development;  
CC premature senescence; increased necrosis; increase in seedling or plant  
CC size; decreased plant size; leaf morphology; seed morphology; seed  
CC biochemistry; increase in root anthocyanins; increase in plant

CC anthocyanins, or alteration in light response or shade avoidance. The  
CC transgenic plant, polynucleotides and polypeptides are useful in  
CC bioinformatic search methods. This sequence represents a plant  
CC transcription factor related polynucleotide.

XX SQ Sequence 795 BP; 236 A; 180 C; 175 G; 204 T; 0 U; 0 Other;

Query Match 7.7%; Score 581; DB 12; Length 795;

Best Local Similarity 100.0%; Pred. No. 3.5e-164;

Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4232 ATGGTGGAGCGTCTCGGATCGCAAAAGTCCCAATTCGCTGAAGCAGCTCTAAATTCGCCCT 4291

Db 1 ATGGTGGAGCGTCTCGGATCGCAAAAGTCCCAATTCGCTGAAGCAGCTCTAAATTCGCCCT 60

QY 4292 AGATGTGACTCAACCAATCTAAGTTCGTGTACTTCAATAAATCTATAGCCCTTACTCAACCT 4351

Db 61 AGATGTGACTCAACCAATCTAAGTTCGTGTACTTCAATAAATCTATAGCCCTTACTCAACCT 120

QY 4352 CGCCATTCTGCAAAACATGTCGCTATTTGACACGTCGGCGGTTCTTTGAGGAATGTT 4411

Db 121 CGCCATTCTGCAAAACATGTCGCTATTTGACACGTCGGCGGTTCTTTGAGGAATGTT 180

QY 4412 CCTGTTGGAGGCGCTTTAGGAGGACAAAGACAAATCCAGATCCGAATCTACGCTC 4471

Db 181 CCTGTTGGAGGCGCTTTAGGAGGACAAAGACAAATCCAGATCCGAATCTACGCTC 240

QY 4472 GTGGTCTCGACTGATTAATCTACTAGTACTTCACTCACTTCTCGCCCAAGTTACTCA 4531

Db 241 GTGGTCTCGACTGATTAATCTACTAGTACTTCACTCACTTCTCGCCCAAGTTACTCA 300

QY 4532 AACCTTAGCAAGTTTCATAGTACGTCGCTCAAAATCCCGAGTTTAATTCCAACTTGCCTATC 4591

Db 301 AACCTTAGCAAGTTTCATAGTACGTCGCTCAAAATCCCGAGTTTAATTCCAACTTGCCTATC 360

QY 4592 TTGCTCTCTCAAGCGCTTCGAGATTACAAATTCAGCAACACTGATTAATTTGGT 4651

Db 361 TTGCTCTCTCAAGCGCTTCGAGATTACAAATTCAGCAACACTGATTAATTTGGT 420

QY 4652 GGAACCTCAATAAGCAACATGATAGTGGTATGAGTTCTAGTGGTGGGATCTTTGATGCA 4711

Db 421 GGAACCTCAATAAGCAACATGATAGTGGTATGAGTTCTAGTGGTGGGATCTTTGATGCA 480

QY 4712 TGGAGTAATCTCCATCAACAAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGA 4771

Db 481 TGGAGTAATCTCCATCAACAAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGA 540

QY 4772 TTGGTCAATCTTCAACCGCTTATATCCATTACTAGAAGG 4812

Db 541 TTGGTCAATCTTCAACCGCTTATATCCATTACTAGAAGG 581

RESULT 15

ABX56844

ID ABX56844 standard; DNA; 577 BP.

XX AC ABX56844;

XX DT 20-FEB-2003 (first entry)

XX DE Arabidopsis thaliana polynucleotide #196.

XX KW Thale cress; gene; ds; genetic manipulation; plant; biosynthesis;

XX KW genetic modification; environmental stress; disease resistance;

XX KW fungicide; insecticide; stress tolerance.

XX OS Arabidopsis thaliana.

XX OS US2002040489-A1.

XX PN 04-APR-2002.

XX PF 26-JAN-2001; 2001US-00770152.

XX

PR

XX

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

PA

27-JAN-2000; 2000US-0178503P.

(GORL/) GORLACH J.

(ANYX/) AN Y.

(HAMI/) HAMILTON C M.

(PRIC/) PRICE J L.

(RAIN/) RAINES T M.

(YUYU/) YU Y.

(RAME/) RAMEAKA J G.

(PAGE/) PAGE A.

(MATH/) MATHW A V.

(LEDE/) LEDFORD B L.

(WOES/) WOESSNER J P.

(HAAS/) HAAS W D.

(GARC/) GARCIA C A.

(KRIC/) KRICKER M.

(SLAT/) SLATER T.

(DAVI/) DAVIS K R.

(ALLE/) ALLEN K.

(HOFF/) HOFFMAN N.

(HURB/) HURBAN P.

XX

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;  
Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
Hurban P;

WPI; 2003-110410/10.

Novel Arabidopsis thaliana nucleic acid useful for identifying homologous  
or related genes, and to create genetically modified and transgenic  
organisms, such as plant cells and plants.

Claim 1; SEQ ID NO 196; 45pp; English.

The invention relates to Arabidopsis thaliana nucleic acid sequences. The  
DNA sequences and the polypeptides they encode are useful for identifying  
homologous or related genes, for producing compositions that modulate the  
expression or function of the polypeptides, for mapping functional  
regions of the protein, in diagnosis, for studying associated  
physiological pathways, for genetic manipulation of cells, preferably  
plant cells, in screening assays of various plant strains to determine  
the strains that are capable of withstanding a particular disease or  
environmental stress, for enhancing or inhibiting production of  
biosynthetic products in plants and to create genetically modified and  
transgenic organisms, such as plant cells and plants. Transgenic plants  
are useful for introducing or improving disease resistance and stress  
tolerance in plants, screening biologically active agents, such as  
fungicides and insecticides, and for elucidating biochemical pathways.  
Sequences ABX56649-ABX57647 represent Arabidopsis thaliana  
polynucleotides of the invention. Note: The sequence data for this patent  
did not form part of the printed specification but was obtained in  
electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)

Sequence 577 BP; 169 A; 80 C; 137 G; 191 T; 0 U; 0 Other;

Query Match 7.6%; Score 577; DB 10; Length 577;

Best Local Similarity 100.0%; Pred. No. 5.9e-163;

Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4926 CAAGAAATGTGAAGCGGGAAGAGATGATCAGGATCGGGTAGGGATGGAGTGA 4985

Db 1 CAAGAAATGTGAAGCGGGAAGAGATGATCAGGATCGGGTAGGGATGGAGTGA 60

QY 4986 ATAACCTATCAAGAACTTTTGGGTATATCAATAACTCAGGACGACGGAAT 5045

Db 61 ATAACCTATCAAGAACTTTTGGGTATATCAATAACTCAGGACGACGGAAT 120

QY 5046 ACACATCATGGGAGGTAACTCTTGGACCGGTTTACCTCCCAACTCAACAGGCC 5105

Db 121 ACACATCATGGGAGGTAACTCTTGGACCGGTTTACCTCCCAACTCAACAGGCC 180

Qy 5106 ATCTCTCATTCTAAGTACTCAGCACTAGCTATTCTTGATGATTCTTTGTTGGTTGGGT 5165  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
181 ATCTCTCATTCTAAGTACTCAGCACTAGCTATTCTTGATGATTCTTTGTTGGTTGGGT 240  
Qy 5166 GTACATTGGTCTTCTCATGCGAGTTATTGCTGAGGAAGATCAAAACCATGCAGCTATATC 5225  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
241 GTACATTGGTCTTCTCATGCGAGTTATTGCTGAGGAAGATCAAAACCATGCAGCTATATC 300  
Qy 5226 CAAAGGCTAAATTTTGAGGCTCAAAGGAAAGGTATGGTTATAAACTATCTTTTGAICTT 5285  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 CAAAGGCTAAATTTTGAGGCTCAAAGGAAAGGTATGGTTATAAACTATCTTTTGAICTT 360  
Qy 5286 TTAAAGATCTTCAAAGTGTGAGTATGTTTATTGGTTGGCTTCTGGTGATATTTATGTTT 5345  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 TTAAAGATCTTCAAAGTGTGAGTATGTTTATTGGTTGGCTTCTGGTGATATTTATGTTT 420  
Qy 5346 TATTAGAATTTGGTCTTATATATTTGGCTATATATAGAGGTGGGTGATATGTATGAATT 5405  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 TATTAGAATTTGGTCTTATATATTTGGCTATATATAGAGGTGGGTGATATGTATGAATT 480  
Qy 5406 CAAGAGTTGATGTTGAAACTTTTTTGTGTGTTTCATTGAATAATCATCGAATTCCTCAATT 5465  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 CAAGAGTTGATGTTGAAACTTTTTTGTGTGTTTCATTGAATAATCATCGAATTCCTCAATT 540  
Qy 5466 TCTTGAGACCCATTATGAGACATTTGAGACATCTATA 5502  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
541 TCTTGAGACCCATTATGAGACATTTGAGACATCTATA 577

Search completed: December 31, 2005, 07:17:46  
Job time : 2705 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2005, 05:09:59 ; Search time 24334 Seconds  
(without alignments)  
17706.639 Million cell updates/sec

Title: US-10-650-249-1  
Perfect score: 7580  
Sequence: 1 agctctataattcaagaga.....gcgcacacgggtgagact 7580

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:  
1: gb\_ba:  
2: gb\_in:  
3: gb\_env:  
4: gb\_om:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pr:  
9: gb\_ro:  
10: gb\_sts:  
11: gb\_sv:  
12: gb\_un:  
13: gb\_vl:  
14: gb\_htg:  
15: gb\_pl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6130	80.9	103240	15	ATT22E16
2	1402	18.5	10138	11	AF187951
3	1402	18.5	10450	11	AF218466
4	708	9.3	1457	15	AK221402
5	703	9.3	1274	15	AF155818
6	368	4.9	2000	6	AX508868
7	362	4.8	2939	15	AB017485
8	332	4.4	1431	15	AJ616346
9	332	4.4	4462	11	VFO551314
10	332	4.4	9663	11	AF408991
11	332	4.4	12537	11	AF788908
12	332	4.4	12607	11	AY737283
13	331	4.4	8903	11	AY234331
14	331	4.4	8911	11	AY234330
15	331	4.4	8958	11	AY234327
16	331	4.4	9146	11	AF458479
17	331	4.4	9154	11	AF458478
18	331	4.4	9393	11	AY234329

C 19	331	4.4	9401	11	AY234328
C 20	331	4.4	9769	11	AY456904
C 21	331	4.4	12739	6	CS007934
C 22	331	4.4	12739	6	CS025573
C 23	301	4.0	301	15	ATH524488
C 24	300	4.0	552	6	AX405124
C 25	299	3.9	323	6	AR643977
C 26	299	3.9	323	6	AX044091
C 27	222	2.9	470	6	AR437667
C 28	222	2.9	470	6	AX026717
C 29	222	2.9	711	6	A05521
C 30	222	2.9	711	11	ARPT100
C 31	222	2.9	5767	6	AR575401
C 32	222	2.9	5767	6	AX275254
C 33	222	2.9	6426	6	BD209862
C 34	222	2.9	6500	6	BD209863
C 35	222	2.9	6750	6	BD209861
C 36	222	2.9	7074	6	BD209860
C 37	222	2.9	10100	6	BD209864
C 38	222	2.9	10100	6	BD209864
C 39	222	2.9	10166	6	BD209867
C 40	222	2.9	10166	6	BD209867
C 41	222	2.9	10240	6	BD209865
C 42	222	2.9	10272	6	BD209866
C 43	222	2.9	14446	6	AR575402
C 44	222	2.9	14446	6	AX275255
C 45	221	2.9	2728	6	A51131

#### ALIGNMENTS

RESULT 1  
ATT22E16  
LOCUS ATT22E16 103240 bp DNA linear PLN 16-APR-2005  
DEFINITION Arabidopsis thaliana DNA chromosome 3, BAC clone T22E16.  
ACCESSION AL132975  
VERSION AL132975.1 GI:6434228  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
REFERENCE 1  
AUTHORS Benes, V., Wurmbach, E., Drzonek, H., Ansoerge, W., Mewes, H. W.,  
Lemcke, K., Mayer, K. F. X., Quetier, F. and Salanoubat, M.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 103240)  
AUTHORS EU Arabidopsis sequencing project.  
TITLE Direct Submission  
JOURNAL Submitted (23-FEB-2000) MIPS, at the Max-Planck-Institut fuer  
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:  
lemckemips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project  
Coordinator: Marcel Salanoubat and Francis Quetier, Groupement  
d'Interet Public, Centre National de Sequencage - GENOSCOPE; 2 rue  
Gaston Cremieux, BP191, 91006 Evry Cedex, France;  
http://www.genoscope.cns.fr  
Information on performance of analysis and a more detailed  
annotation of this entry and other sequences of chromosomes 3, 4  
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/chal/.  
Location/Qualifiers  
1..103240  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/ecotype="Columbia"  
1..8852  
/note="overlap to BAC T26112 (EMBL:AL132954)"  
992..3164  
/gene="T22E16.10"  
Join(992..1181,2134..3164)  
misc\_feature  
gene  
CDS



/gene="T22E16.10"  
 /note="similarity to predicted proteins, Arabidopsis thaliana"  
 /codon\_start=1  
 /product="putative protein"  
 /protein\_id="CAB75893.1"  
 /db\_xref="GI:7076778"  
 /db\_xref="UniProt/TREMBL:Q9M2U3"  
 /translation="MGPIKTIKKKRAKKVDRNVLLAATAAATSASAAAAANNDDDD  
 DSSQSLDWDGFRRIYGGSTDPKTFESVFKIRKTFDYICSLVKADFTAKPANFS  
 DNGSPSLNDRLVALRGLSGESLSVIGETFGMNSTVQITWRVSEMEERAIHH  
 LSWPKLDEIKSKFEKISGLPNCCGAIIDTHVMNLPAPSPNKMVLDCGNKFSMTLQ  
 AVVDPMDRLDIAGWPGSLDDVVLKNSGFYKLVKKGRLNGEKLPLSETELREYI  
 VDSGFFLLPWLITPYQKPTSLPQTEFNKHSSEATKAQWALSCLKDRWRLLNGVM  
 MPDRNRLLPIIFVCLLHNIIIDMEDQTLDDQFLSQDHDMYRQRSCKLADASSVLR  
 DELSPQLCGKNSSA"  
 992. .1181  
 /gene="T22E16.10"  
 /number=1  
 /number=1  
 1182. .2133  
 /gene="T22E16.10"  
 /number=1  
 2134. .3164  
 /gene="T22E16.10"  
 /number=2  
 3660. .5330  
 /gene="T22E16.20"  
 complement (3660. .3728,3821. .4575,5038. .5077,  
 5262. .5330)  
 /gene="T22E16.20"  
 complement (3660. .3728,3821. .4575,5038. .5077,  
 5262. .5330)  
 /gene="T22E16.20"  
 /note="strong similarity to synaptic glycoprotein SC2  
 spliced variant, Homo sapiens, EMBL:AF038958"  
 /codon\_start=1  
 /product="glycoprotein-like"  
 /protein\_id="CAB75894.1"  
 /db\_xref="GI:7076779"  
 /db\_xref="GOA:Q9M2U2"  
 /db\_xref="UniProt/TREMBL:Q9M2U2"  
 /translation="MKVTVVSRSREVLKAPLDLPDSATVADLQAFHAKRKKFYPSR  
 QRLTLVPTGSKDKPVLNSKSKLKEYCDGNNSLTVPFKDLGAQVSYRTLFFFEYLG  
 PLLIYVFPYFVFLGSGEDCVIHPVQTIYAMTWCFYFKRILETFVFRFSHATS  
 PTGNVFNCAIYWSFGAYIAYVNHPLVTPVSDLQMKIGFGFLVCQVANFYCHILLK  
 NLRDPSGAGGYQIPRGFLFNI VTCANYTTEIYQWLGNFIATQTIAGYVFLAVAALIMT  
 NNALGKHSRLRKIFDGKDGKPKYPRRWILPPFL"  
 complement (3660. .3728)  
 /gene="T22E16.20"  
 /number=1  
 complement (3729. .3820)  
 /gene="T22E16.20"  
 /number=1  
 complement (3821. .4575)  
 /gene="T22E16.20"  
 /number=2  
 complement (4576. .5037)  
 /gene="T22E16.20"  
 /number=2  
 complement (5038. .5077)  
 /gene="T22E16.20"  
 /number=3  
 complement (5078. .5261)  
 /gene="T22E16.20"  
 /number=3  
 complement (5262. .5330)  
 /gene="T22E16.20"  
 /number=4  
 10334. .11350  
 /gene="T22E16.30"  
 join (10334. .11042,11136. .11350)  
 /gene="T22E16.30"

/note="EMBL:AF155818 most likely is not full length,  
 therefore N-terminal stretch"  
 /codon\_start=1  
 /product="zinc finger protein OBP3"  
 /protein\_id="CAB75895.1"  
 /db\_xref="GI:7076780"  
 /db\_xref="GOA:Q9M2U1"  
 /db\_xref="UniProt/Swiss-Prot:Q9M2U1"  
 /translation="WQGNQHQLECVTTDDQNNYLRQLSSPPTSQVAGSSQARVNSM  
 VERARIKVPLEALNCPDCSTNTKFCYNNYSLTQPRHCKTCRVYWTGSSLRN  
 VFGGKFRNRKRSKRSKSTVVVSDNTDITSTSLSRFSYSPKSFHYGQIPERNSN  
 LPILPLSLGDYNSNTGLDPFGGTQISNMISGMSGGILDAWRIPSPQQAQPPFL  
 INTTGLQSNALYPLLEGVSATQTRNVKAEENDQDRGDGYNLNRNPLGNINI  
 NSGRNEEYTSWGSNSWTGFTSNNSTGHLSP"  
 10334. .11042  
 /gene="T22E16.30"  
 /number=1  
 11043. .11135  
 /gene="T22E16.30"  
 /number=1  
 11136. .11350  
 /gene="T22E16.30"  
 /number=2  
 13994. .15369  
 /gene="T22E16.40"  
 join (13994. .14039,14159. .14261,14513. .14610,14695. .14795,  
 15214. .15369)  
 /gene="T22E16.40"  
 /note="Contains Ubiquitin-conjugating enzymes active site  
 AA79-93"  
 /codon\_start=1  
 /product="ubiquitin-conjugating enzyme UBC3"  
 /protein\_id="CAB75896.1"  
 /db\_xref="GI:7076781"  
 /db\_xref="GOA:P42747"  
 /db\_xref="UniProt/Swiss-Prot:P42747"  
 /translation="MANNOASLLQKLDCKPVDGFSAGLVDEKNVFWQSVSIMG  
 PPDTLVEGGFNAIMSPENYVPSPPTFTSEMHWPVNSDGKVCISLTHPKPGDDPH  
 GYELASERWTPVHTVESIVLSIISMLSGPNDESPANVEAAKEWRDNRRAEFRKKVSRVC  
 RRSQEML"  
 13994. .14039  
 /gene="T22E16.40"  
 /number=1  
 14040. .14158  
 /gene="T22E16.40"  
 /number=1  
 14159. .14261  
 /gene="T22E16.40"  
 /number=2  
 14262. .14512  
 /gene="T22E16.40"  
 /number=2  
 14513. .14610  
 /gene="T22E16.40"  
 /number=3  
 14611. .14694  
 /gene="T22E16.40"  
 /number=3  
 14695. .14795  
 /gene="T22E16.40"  
 /number=4  
 14796. .15213  
 /gene="T22E16.40"  
 /number=4  
 15214. .15369  
 /gene="T22E16.40"  
 /number=5  
 17259. .18094  
 /gene="T22E16.50"  
 join (17259. .17479,17564. .17699,17867. .18094)  
 /gene="T22E16.50"  
 /note="similarity to MtN24, Medicago truncatula,"

EMBL:MTY15290  
Contains Prokaryotic membrane lipoprotein lipid attachment  
site AA164-174  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAB75897.1"  
/db\_xref="GI:7076782"  
/db\_xref="InterPro:IPR006459"  
/db\_xref="UniProt/TREMBL:Q9M2U0"  
/translation="MRSPHFRNGESPTRLDTHPHSTVTAOKLRRFNSLILLRLAS  
FSPSLASAVFMLTNSRGSASPHWYDFDAERFVFVANAIVALYSPVEMGCVNEFSRET  
TLNPEAPQWFDFGHQDFVSFLLLSAGSAAALARPMRGDDICTANKAFCLSDVAIG  
LGFPAELFLAFSSCFSGFRVACFLITGRFHLYS"  
17259..17479  
exon

Query Match	80.9%	Score 6130;	DB 15;	Length 103240;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 6130;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	AGCTCTATTAAATTC	AAGAGAGCAGCAAAATAAAGCAAAAACTCAAAACCTTAAGTTTCTCGA	60	
DB 6232	AGCTCTATTAAATTC	AAGAGAGCAGCAAAATAAAGCAAAAACTCAAAACCTTAAGTTTCTCGA	6291	
QY 61	ATATGAAGGGGTAGAT	TAATCAATCTCTCAACTAGTTAAAAAAGTAATGATAAATA 120		
DB 6292	ATATGAAGGGGTAGAT	TAATCAATCTCTCAACTAGTTAAAAAAGTAATGATAAATA 6351		
QY 121	AAAAACAATGGATCAAT	TAAGAGACAGTAGTTTATGATATATATGTTGGGATCGATTA 180		
DB 6352	AAAAACAATGGATCAAT	TAAGAGACAGTAGTTTATGATATATATGTTGGGATCGATTA 6411		
QY 181	GTTGACATCACAAAGAT	CAAAATAATGACCGGTAAATGCCCAATAACCAAGGCGCAAT 240		
DB 6412	GTTGACATCACAAAGAT	CAAAATAATGACCGGTAAATGCCCAATAACCAAGGCGCAAT 6471		
QY 241	CATGCGATATTCAAACACCT	TAATGTCATTCAAATCTAATACTAATACTATCCTCAATTA 300		
DB 6472	CATGCGATATTCAAACACCT	TAATGTCATTCAAATCTAATACTAATACTATCCTCAATTA 6531		
QY 301	CAAAAGCTAGTAGATTTT	CTCAAGTGAGCAACAGTCTAATCTCTCTGAAAAAATCTGTT 360		
DB 6532	CAAAAGCTAGTAGATTTT	CTCAAGTGAGCAACAGTCTAATCTCTCTGAAAAAATCTGTT 6591		
QY 361	TTTTCCGAGTGAAATTA	TCCATCTACTACTACAAATTTGAGCAATTAATCTAGTTTCT 420		
DB 6592	TTTTCCGAGTGAAATTA	TCCATCTACTACTACAAATTTGAGCAATTAATCTAGTTTCT 6651		
QY 421	CTGCAACTTTAAACCAAGT	GTAACAAGTCAACACAGATCTAGCGTAAACACCTTAAG 480		
DB 6652	CTGCAACTTTAAACCAAGT	GTAACAAGTCAACACAGATCTAGCGTAAACACCTTAAG 6711		
QY 481	CTAGTACTTAATAGATTA	TGCTATTTTTCGACCAATTAATTAATTAATCTTCCAGCCT 540		
DB 6712	CTAGTACTTAATAGATTA	TGCTATTTTTCGACCAATTAATTAATTAATCTTCCAGCCT 6771		
QY 541	TTTCGTAAGAAATATG	CGCATGTTCTTGTGGAATCTAGGAATCTTCTTCTACTAAGATTG 600		
DB 6772	TTTCGTAAGAAATATG	CGCATGTTCTTGTGGAATCTAGGAATCTTCTTCTACTAAGATTG 6831		
QY 601	GCATGCACTGTAAGATTT	CCATTTGATATCTCGATATTTTCCACCTTAATAAT 660		
DB 6832	GCATGCACTGTAAGATTT	CCATTTGATATCTCGATATTTTCCACCTTAATAAT 6891		
QY 661	CTTGAAATTAAGATCA	AAATTAATGAGAACATATATTTGATCAATGTTTGAATAGCCT 720		
DB 6892	CTTGAAATTAAGATCA	AAATTAATGAGAACATATATTTGATCAATGTTTGAATAGCCT 6951		
QY 721	TATATAGTGTAGATGT	GAACTATGGAACAACAAATAGGAAAGAAAGCTGCGAG 780		
DB 6952	TATATAGTGTAGATGT	GAACTATGGAACAACAAATAGGAAAGAAAGCTGCGAG 7011		
QY 781	AGAAAGGTGTGCTTTA	CAACTATGAATATGGGCTCATGAGATGTACACTACAGTAAGCG 840		

DB 7012	AGAAGGTGTGCTTTT	TACAACATGAAATATGGGCTCATGAGATGTACACTACAGATAAGCG 7071		
QY 841	AAAGATTCCTGTTG	CGATGAAATGTTTACTAATAAAAAAACAACATGCGCATTTCTATAA 900		
DB 7072	AGATTCCTGTTG	CGATGAAATGTTTACTAATAAAAAAACAACATGCGCATTTCTATAA 7131		
QY 901	AGACGAATTCCTTT	TAAATAAATTTCTATAAACAATAAAGATAGTCTCTCTTAA 960		
DB 7132	AGACGAATTCCTTT	TAAATAAATTTCTATAAACAATAAAGATAGTCTCTCTTAA 7191		
QY 961	ACATCGCAAAAGAA	TATATAGATTTACCGTATCAGATTTTTCATCAATTTTATATTTT 1020		
DB 7192	ACATCGCAAAAGAA	TATATAGATTTACCGTATCAGATTTTTCATCAATTTTATATTTT 7251		
QY 1021	TGAGCTTTGAAAGAT	TAAACATGATCTGTCGTGTCCTGCTATTTACCCCTAGA 1080		
DB 7252	TGAGCTTTGAAAGAT	TAAACATGATCTGTCGTGTCCTGCTATTTACCCCTAGA 7311		
QY 1081	AGAAGTGAACATG	GAACTTTTATGATTTGCAATAACGCGAGCTAGCTTCTCCCTACTTG 1140		
DB 7312	AGAAGTGAACATG	GAACTTTTATGATTTGCAATAACGCGAGCTAGCTTCTCCCTACTTG 7371		
QY 1141	TCCAATAGATGA	GACATTTATCACTCAGGTCAGCTTTCGAAGCGCAACATATCGACA 1200		
DB 7372	TCCAATAGATGA	GACATTTATCACTCAGGTCAGCTTTCGAAGCGCAACATATCGACA 7431		
QY 1201	AAAAATCGTTT	TAGCTCTATCATCTGTCCTTTTGAAGAAAAATATCAACATATCAAAATACA 1260		
DB 7432	AAAAATCGTTT	TAGCTCTATCATCTGTCCTTTTGAAGAAAAATATCAACATATCAAAATACA 7491		
QY 1261	TATACACACTCCC	CAATATATAACCAATATATATGTTTACTACGAAATTCAT 1320		
DB 7492	TATACACACTCCC	CAATATATAACCAATATATATGTTTACTACGAAATTCAT 7551		
QY 1321	GATATGCTCTTT	TGAACCAACTGTAATCAAGCTTAAGCAATATATCTATTAT 1380		
DB 7552	GATATGCTCTTT	TGAACCAACTGTAATCAAGCTTAAGCAATATATCTATTAT 7611		
QY 1381	ATTCCTTTCTT	TGTCCTCAACTCATCCCTTTCTAAGTAATTTCAAAATTAATTTGGAAGTTT 1440		
DB 7612	ATTCCTTTCTT	TGTCCTCAACTCATCCCTTTCTAAGTAATTTCAAAATTAATTTGGAAGTTT 7671		
QY 1441	TGTCATCTAAAT	TGAAGTTCTTTTACCGGATCATTTTGTCTCAGGTGATATATCTTT 1500		
DB 7672	TGTCATCTAAAT	TGAAGTTCTTTTACCGGATCATTTTGTCTCAGGTGATATATCTTT 7731		
QY 1501	CTTAGTCTGATA	CGAAAAAATTAATGATTAAGAGAGACATGTTTGATGTTA 1560		
DB 7732	CTTAGTCTGATA	CGAAAAAATTAATGATTAAGAGAGACATGTTTGATGTTA 7791		
QY 1561	TATTTTCTG	TGTAAGAAATCATCTGATTAAGAACTATATAGTTAGGGATATGTTTCATGT 1620		
DB 7792	TATTTTCTG	TGTAAGAAATCATCTGATTAAGAACTATATAGTTAGGGATATGTTTCATGT 7851		
QY 1621	TGTTGTG	CGACATAGTGGGTCATTTAAAGAAGTCTGATTTAAATTTTACGTTTGGAC 1680		
DB 7852	TGTTGTG	CGACATAGTGGGTCATTTAAAGAAGTCTGATTTAAATTTTACGTTTGGAC 7911		
QY 1681	CACAAATCTT	CTTTTAGAAATCGGGAGCTGGGACCTTCTCAACATGTCCTCTTT 1740		
DB 7912	CACAAATCTT	CTTTTAGAAATCGGGAGCTGGGACCTTCTCAACATGTCCTCTTT 7971		
QY 1741	ACTAATCTT	TACGTACCCCTCACATTCGTAACCATAAAAATCAATAATAATATAGAGAC 1800		
DB 7972	ACTAATCTT	TACGTACCCCTCACATTCGTAACCATAAAAATCAATAATAATATAGAGAC 8031		
QY 1801	TGGTATCAT	ATATTCGAAATATTTTTCATTAATCAATGTTTATCGGTAAAGTTATATAGG 1860		
DB 8032	TGGTATCAT	ATATTCGAAATATTTTTCATTAATCAATGTTTATCGGTAAAGTTATATAGG 8091		
QY 1861	GGTATACAT	CAAGAAATCAGGAAAGATTAATAAACCACTTGTCTGAAATCATGATTTGATT 1920		



4081 AATTAGTACAAATTAATGTTAATATGACGAGGAAACCAACATCAGCTAGAAATGTTGCAC 4140  
10312 AATTAGTACAAATTAATGTTAATATGACGAGGAAACCAACATCAGCTAGAAATGTTGCAC 10371  
4141 AACTGACCAAGAACCCCTAAATTAATTAATGAGGAGCTCTCATACACCAAGCTTCTCAGGT 4200  
10372 AACTGACCAAGAACCCCTAAATTAATTAATGAGGAGCTCTCATACACCAAGCTTCTCAGGT 10431  
4201 TGCAGGTTCCAGTCAAGCTAGAGTGAATCAATGTTGGAAACGTTCTCGATCGCAAAAGT 4260  
10432 TGCAGGTTCCAGTCAAGCTAGAGTGAATCAATGTTGGAAACGTTCTCGATCGCAAAAGT 10491  
4261 CCCATTGCCCTGAAGCAGCTCTAAATTTGCCCTAGATGTAACCAACAACTAAGTTCTG 4320  
10492 CCCATTGCCCTGAAGCAGCTCTAAATTTGCCCTAGATGTAACCAACAACTAAGTTCTG 10551  
4321 TTACTTTCAATAACTATAGCTTACTCAACCTGCCATTTCTGCAAAACATGTCGTCGCTA 4380  
10552 TTACTTTCAATAACTATAGCTTACTCAACCTGCCATTTCTGCAAAACATGTCGTCGCTA 10611  
4381 TTGGACACGTTGGCGGTTCTTGGAGGAATGTTCTGTTGGAGGAGGCTTTAGGAGGAACAA 4440  
10612 TTGGACACGTTGGCGGTTCTTGGAGGAATGTTCTGTTGGAGGAGGCTTTAGGAGGAACAA 10671  
4441 GAGAGCAATCCAGATCGAAATCTACGGTCTGTTGCTCGACTGATAAATACTACTAGTAC 4500  
10672 GAGAGCAATCCAGATCGAAATCTACGGTCTGTTGCTCGACTGATAAATACTACTAGTAC 10731  
4501 TTCAATCACTTCTTCCGCAAGTTACTCAAAACCTCAGCAAGTTTCATAGCTACCGTCA 4560  
10732 TTCAATCACTTCTTCCGCAAGTTACTCAAAACCTCAGCAAGTTTCATAGCTACCGTCA 10791  
4561 AATCCCGGAGTTTAAATTCAACTTGGCCATCTTGGCTCTCTCCAAAGCCTTGGAGATTA 4620  
10792 AATCCCGGAGTTTAAATTCBACTTGGCCATCTTGGCTCTCTCCAAAGCCTTGGAGATTA 10851  
4621 CAATTCAGAACCACTGGATTAATTTTGGTGAATCTCAATTAAGCAACATGATAAGTGG 4680  
10852 CAATTCAGAACCACTGGATTAATTTTGGTGAATCTCAATTAAGCAACATGATAAGTGG 10911  
4681 TATGAGTTCTAGTGGTGGATCTTGGATGATGAGAAATACCTCCATCAACCAAGCTCA 4740  
10912 TATGAGTTCTAGTGGTGGATCTTGGATGATGAGAAATACCTCCATCAACCAAGCTCA 10971  
4741 GCAATTCCTCTTCTGATCAACACTACCGGATTTGGTGAATCTTCAAAACGCTTATATCC 4800  
10972 GCAATTCCTCTTCTGATCAACACTACCGGATTTGGTGAATCTTCAAAACGCTTATATCC 11031  
4801 ATTTACTAGAGGTAAGGAGGTTTAAATCAAGGTGATTTCTCAACAGAGAGTATGATTA 4860  
11032 ATTTACTAGAGGTAAGGAGGTTTAAATCAAGGTGATTTCTCAACAGAGAGTATGATTA 11091  
4861 TTCCAATCAGCTAATGTTTAAAGCCCTTATGAGATTTTCTTCAAGCGGGTTAGCCGCTAC 4920  
11092 TTCCAATCAGCTAATGTTTAAAGCCCTTATGAGATTTTCTTCAAGCGGGTTAGCCGCTAC 11151  
4921 GCAAAACAGAAATGTCAGGCGGAGAGATGATCAGGATCGGGGTAGGATCGGGATGG 4980  
11152 GCAAAACAGAAATGTCAGGCGGAGAGATGATCAGGATCGGGGTAGGATCGGGATGG 11211  
4981 AGTGAATAACTTATCAAGAAACCTTTTGGGTAAATATCAACATAAACTCAGGCGAGGAACGA 5040  
11212 AGTGAATAACTTATCAAGAAACCTTTTGGGTAAATATCAACATAAACTCAGGCGAGGAACGA 11271  
5041 GGAATAACATCATGAGGAGGTAACAGTTCTTGGACCGGTTTCCACCTCCAACTCAAC 5100  
11272 GGAATAACATCATGAGGAGGTAACAGTTCTTGGACCGGTTTCCACCTCCAACTCAAC 11331  
5101 AGGCCATCTCTCAATCTAAGTACTCAGCACTAGCTATCTTGTGATGATCTTTTGTGGTT 5160  
11332 AGGCCATCTCTCAATCTAAGTACTCAGCACTAGCTATCTTGTGATGATCTTTTGTGGTT 11391  
5161 GGGGTGTACATTGGTGTCTGTCATGCGAGTTATTGCTGAGGAAGATCAAAACCATGCGCT 5220

11392 GGGGTGTACATTGGTGTCTTGTCAATGCGAGTTATTGCTGAGGAAGATCAAAACCATGCGCT 11451  
5221 ATATCCAAAGGCTAAATTTTGGGCTCAAAAGGAAGGTATGGTTTATAAACTATCTTTTGTG 5280  
11452 ATATCCAAAGGCTAAATTTTGGGCTCAAAAGGAAGGTATGGTTTATAAACTATCTTTTGTG 11511  
5281 ATCTTTTAAAGATCTTCAAGTGTGAGTATGTTTATTTGGTGGCTCTGGTGATATTTA 5340  
11512 ATCTTTTAAAGATCTTCAAGTGTGAGTATGTTTATTTGGTGGCTCTGGTGATATTTA 11571  
5341 TGTTTTATTAGAAATTTGGTCTTATATATTGGCTTATATATAGAGTCTGGGTGATATGTAT 5400  
11572 TGTTTTATTAGAAATTTGGTCTTATATATTGGCTTATATATAGAGTCTGGGTGATATGTAT 11631  
5401 GAATTCAGAGTGTATGTTTGGAAACCTTTTGTGTGTTTCATTGAATTAATCATCGAATTTCT 5460  
11632 GAATTCAGAGTGTATGTTTGGAAACCTTTTGTGTGTTTCATTGAATTAATCATCGAATTTCT 11691  
5461 CAATTTCTTGGAGACCCATTTATGAGACATTTGAGACATCTATAGAACATATATGTAATGTA 5520  
11692 CAATTTCTTGGAGACCCATTTATGAGACATTTGAGACATCTATAGAACATATATGTAATGTA 11751  
5521 TATTAACCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 5580  
11752 TATTAACCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 11811  
5581 TGTAAATATCGAGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 5640  
11812 TGTAAATATCGAGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 11871  
5641 CTATATATCTTATTTCTGATTTCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5700  
11872 CTATATATCTTATTTCTGATTTCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 11931  
5701 GACACAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAAC 5760  
11932 GACACAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAAC 11991  
5761 GAGAATTTAGCACTTAAATTTAGCACTTAAATTTAGCACTTAAATTTAGCACTTAAATTTAGCACT 5820  
11992 GAGAATTTAGCACTTAAATTTAGCACTTAAATTTAGCACTTAAATTTAGCACTTAAATTTAGCACT 12051  
5821 CGATTTAAATTTGTGTTCAATATGATGTTCAATATGATGTTCAATATGATGTTCAATATGATGTT 5880  
12052 CGATTTAAATTTGTGTTCAATATGATGTTCAATATGATGTTCAATATGATGTTCAATATGATGTT 12111  
5881 TGACTGAAACCAATGCAATATCTCTCAAGTCTCAACCTATGAAAGATCATGTAACCAATAG 5940  
12112 TGACTGAAACCAATGCAATATCTCTCAAGTCTCAACCTATGAAAGATCATGTAACCAATAG 12171  
5941 ACTATCATCATGATTAATGTTAAATGTCATGATCTATAATGTTAAATGTTAAATGTTAAATGTT 6000  
12172 ACTATCATCATGATTAATGTTAAATGTCATGATCTATAATGTTAAATGTTAAATGTTAAATGTT 12231  
6001 ATTTATCTGGATATAAGATGCGGTTTAAACCTTCTTCAATTTTCAATTTTGTGTTTATCTTTCT 6060  
12232 ATTTATCTGGATATAAGATGCGGTTTAAACCTTCTTCAATTTTCAATTTTGTGTTTATCTTTCT 12291  
6061 TCTAATACATATGATCAATACACTTTTGTGTTTAAAGAAATTTAAACCTTATTTTCAAC 6120  
12292 TCTAATACATATGATCAATACACTTTTGTGTTTAAAGAAATTTAAACCTTATTTTCAAC 12351  
6121 ATCGATCACA 6130  
12352 ATCGATCACA 12361

RESULT 2

AF187951/c

LOCUS

DEFINITION

ACCESSION

AF187951 10138 bp DNA linear SYN 11-MAY-2000  
Activation-tagging vector pSKI015, complete sequence.

VERSION	AF187951.1	GI:6537289	
KEYWORDS	Activation-tagging vector pSKI015		
SOURCE	Activation-tagging vector pSKI015		
ORGANISM	other sequences; artificial sequences; vectors.		
REFERENCE	1 (bases 1 to 10138)		
AUTHORS	Weigel,D., Ahn,J.H., Blazquez,M.A., Borevitz,J.O., Christensen,S.K., Fanhauser,C., Ferrandiz,C., Kardailsky,I., Malancharuvill,E.J., Neff,M.M., Nguyen,J.T., Sato,S., Wang,Z., Xia,Y., Dixon,R.A., Harrison,M.J., Lamb,C.J., Yanofsky,M.P. and Chory,J.		
TITLE	Activation tagging in Arabidopsis		
JOURNAL	Plant Physiol. 122 (4), 1003-1013 (2000)		
PUBLISHED	10759496		
REFERENCE	2 (bases 1 to 10138)		
AUTHORS	Kardailsky,I. and Weigel,D.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-SEP-1999) Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
FEATURES	Location/Qualifiers		
source	1..10138		
	/organism="Activation-tagging vector pSKI015"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:109189"		
	12..31		
misc_feature	/notes="T7 RNA polymerase promoter"		
misc_feature	36..2204		
	/notes="derived from pUC19"		
misc_feature	2205..2662		
	/notes="derived from phage f1"		
misc_feature	2663..2829		
	/notes="derived from pUC19"		
misc_feature	2830..2849		
	/notes="T7 RNA polymerase promoter"		
misc_feature	2894..4178		
	/notes="cauliflower mosaic virus 35S gene promoter enhancer sequences (tetramer)"		
misc_feature	4270..4294		
	/notes="T-DNA right border"		
misc_feature	7690..7713		
	/notes="T-DNA left border"		
mRNA	complement (8054..8784)		
	/product="octopine synthase"		
	/notes="3', end"		
mRNA	complement (<8785..>9420)		
CDS	complement (8791..9342)		
	/notes="promotes resistance to glutamine synthetase inhibitors"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="AAF15587.1"		
	/db_xref="GI:6537290"		
	/translation="MSPERPADIRREADMPVCTIVNHVYLETSTVNRTEPBPQ EWTDLVLRERYPLWVAEVDGEVAGIAYAGPKARNADWTAEISTVYSPRHQFGL GSTLYTHLLKSLGAQGFKSVAVIGLPNDFSVRHEALGYAPRGMURAAGFKHGNWHD VGFQQLDFSLPVPFPVLPVTEN"		
mRNA	complement (9421..10119)		
	/product="mannopine synthase"		
	/note="5', end and promoter"		
ORIGIN	Query Match 18.5%; Score 1402; DB 11; Length 10138;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 1402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	6179	GTTCGGATCTAGATATACATCAATCCACTTCGTTTGAAGACGTGTTGGACGCTTTC	6238
Db	4259	GTTCGGATCTAGATATACATCAATCCACTTCGTTTGAAGACGTGTTGGACGCTTTC	4200
Qy	6239	TTTTCACCATGTTCTCGTGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGC	6298
Db	4199	TTTTTCCAGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGACCACTGTCGGTAGAGGC	4140
Qy	6299	ATCTTGAAAGATAGACCTTTCTTTATCGCAATGATGGCAATTTGTAGAACCACTCTTCCTT	6358
Db	4139	ATCTTGAAAGATAGACCTTTCTTTATCGCAATGATGGCAATTTGTAGAACCACTCTTCCTT	4080
Qy	6359	TTTCTACTGTCCTTTTCGATGAAGTGAAGTGTGGGCAATGGAATCCGAGAGGTTTCC	6418
Db	4079	TTTCTACTGTCCTTTTCGATGAAGTGAAGTGTGGGCAATGGAATCCGAGAGGTTTCC	4020
Qy	6419	CGATATTACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACGTGATCTTT	6478
Db	4019	CGATATTACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACGTGATCTTT	3960
Qy	6479	GATATTCTTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGGATCTAGATATCA	6538
Db	3959	GATATTCTTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGGATCTAGATATCA	3900
Qy	6539	TCAATCCACTCTCTTTTGAAGACGTGGTTGGAACGCTTCTTTTCCACGATGTCCTCGT	6598
Db	3899	TCAATCCACTCTCTTTTGAAGACGTGGTTGGAACGCTTCTTTTCCACGATGTCCTCGT	3840
Qy	6599	GGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCCTTCC	6658
Db	3839	GGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGATAGCCCTTCC	3780
Qy	6659	TTTATCGCAATGATGSCATTTCTAGAACCATCTTCTCTTCTACTGCTCTTTCCGATGAA	6718
Db	3779	TTTATCGCAATGATGSCATTTCTAGAACCATCTTCTCTTCTACTGCTCTTTCCGATGAA	3720
Qy	6719	GTGACAGATAGCTGGGCAATCGAATCCGAGAGGTTTCCGATATTAACCTTTGTTGAAA	6778
Db	3719	GTGACAGATAGCTGGGCAATCGAATCCGAGAGGTTTCCGATATTAACCTTTGTTGAAA	3660
Qy	6779	AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACG	6838
Db	3659	AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACG	3600
Qy	6839	AGTGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTTGTCTTGAAGA	6898
Db	3599	AGTGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCCACTTGTCTTGAAGA	3540
Qy	6899	CGTGTGTGAACGTCTTCTTTTCCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGG	6958
Db	3539	CGTGTGTGAACGTCTTCTTTTCCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGG	3480
Qy	6959	ACCACGTGCGGTAGAGGCATCTTGAACGATAGCTTTCTTTATCGCAATGATGGCATTT	7018
Db	3479	ACCACGTGCGGTAGAGGCATCTTGAACGATAGCTTTCTTTATCGCAATGATGGCATTT	3420
Qy	7019	GTAGAACCATCTTCTCTTTTCTACTGCTCTTTTCGATGAAGTGAACATAGCTGGGCAATG	7078
Db	3419	GTAGAACCATCTTCTCTTTTCTACTGCTCTTTTCGATGAAGTGAACATAGCTGGGCAATG	3360
Qy	7079	GAATCCGAGAGGTTTCCCGATATTAACCTTTTGTGAAAGTCTCAATAGCCCTCTGGTC	7138
Db	3359	GAATCCGAGAGGTTTCCCGATATTAACCTTTTGTGAAAGTCTCAATAGCCCTCTGGTC	3300
Qy	7139	TTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGTGTCTGCTCCACCATGTT	7198
Db	3299	TTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGTGTCTGCTCCACCATGTT	3240
Qy	7199	GGGGATCTAGATATCAATCAATCCACTTGTGTTGAAGACGTGTTGGAACTGTTCTTTT	7258
Db	3239	GGGGATCTAGATATCAATCAATCCACTTGTGTTGAAGACGTGTTGGAACTGTTCTTTT	3180
Qy	7259	TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATC	7318
Db	3179	TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATC	3120
Qy	7319	TTGAAACGATAGCCCTTTCTTTATCGCAATGATGGCAATTTGTAGAACCACTCTTCCTTTC	7378
Db	3119	TTGAAACGATAGCCCTTTCTTTATCGCAATGATGGCAATTTGTAGAACCACTCTTCCTTTC	3060

QY 7379 TACTGTCTCTTCGATGAAGTACAGATAGCTGGGCAATCGAATCCGAGGAGTTTCCCGA 7438  
Db 3059 TACTGTCTCTTCGATGAAGTACAGATAGCTGGGCAATCGAATCCGAGGAGTTTCCCGA 3000  
QY 7439 TATTACCCCTTTGTTGAAAGTCTCAATAGACCTCTGCTCTCTGAGACTGTATCTTTGAT 7498  
Db 2999 TATTACCCCTTTGTTGAAAGTCTCAATAGACCTCTGCTCTCTGAGACTGTATCTTTGAT 2940  
QY 7499 ATTCTTTGGAGTAGACAGAGAGTGTCTGCTCCACCATGTTGGGGATCCACTAGTCTTAGAG 7558  
Db 2939 ATTCTTTGGAGTAGACAGAGAGTGTCTGCTCCACCATGTTGGGGATCCACTAGTCTTAGAG 2880  
QY 7559 CGGCGCCACCGCGGTGGAGCT 7580  
Db 2879 CGGCGCCACCGCGGTGGAGCT 2858

RESULT 3  
AF218466/c  
LOCUS AF218466 10450 bp DNA circular SYN 11-MAY-2000  
DEFINITION Activation tagging vector pSKI074, complete sequence.  
ACCESSION AF218466  
VERSION AF218466.1 GI:6715465  
KEYWORDS  
SOURCE Activation tagging vector pSKI074  
ORGANISM Activation tagging vector pSKI074  
other sequences; artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 10450)  
AUTHORS Weigel,D., Ahn,J.H., Blazquez,M.A., Borevitz,J.O.,  
Christensen,S.K., Fankhauser,C., Ferrandiz,C., Kardailsky,I.,  
Malancharuvil,E.J., Neff,M.M., Nguyen,J.T., Sato,S., Wang,Z.,  
Xia,Y., Dixon,R.A., Harrison,M.J., Lamb,C.J., Yanofsky,M.F. and  
Chory,J.

TITLE Activation tagging in Arabidopsis  
JOURNAL Plant Physiol. 122 (4), 1003-1013 (2000)  
PUBMED 10759496  
REFERENCE 2 (bases 1 to 10450)  
AUTHORS Weigel,D., Blazquez,M.A., Borevitz,J., Christensen,S.K.,  
Fankhauser,C., Ferrandiz,C., Malancharuvil,E.J., Neff,M.M.,  
Nguyen,J.T., Sato,S., Xia,Y., Wang,Z., Dixon,R.A., Harrison,M.J.,  
Lamb,C.J., Yanofsky,M.F. and Chory,J.  
Direct Submission  
Submitted (22-DEC-1999) Plant Biology Laboratory, The Salk  
Institute for Biological Studies, 10010 N. Torrey Pines Road, La  
Jolla, CA 92037, USA

FEATURES  
Source  
1..10450  
/organism="Activation tagging vector pSKI074"  
/mol\_type="other DNA"  
/db\_xref="taxon:112458"  
12..31  
/note="from T7 RNA polymerase"  
36..2204  
/note="derived from pUC19"  
1862..2205  
/note="derived from phage f1"  
2663..2829  
/note="derived from pUC19"  
2830..2849  
/note="from T7 RNA polymerase"  
2894..4178  
/note="from cauliflower mosaic virus 35S gene promoter"  
4270..4294  
/note="T-DNA right border"  
8054..8784  
/note="T-DNA left border"  
8064..8749  
/note="includes mannopine synthase promoter and 5' UTR  
sequence"  
8756..9739  
/product="neomycin phosphotransferase"  
8771..9565  
/function="confers kanamycin resistance"

/codon\_start=1  
/product="neomycin phosphotransferase"  
/protein\_id="AAF26422.1"  
/db\_xref="GI:6715466"  
/translation="MIBODGLHAGSPAAWVERLFGYDWAQQTIGCSDDAAVFRLSAQR  
PVLVFKTDLGALNELQDEAARLSWLAATTGVPCAAVLDVVTRAGRDMLLGVEVPGDGL  
LSSHLAPAEKVISIMADAMRRRLHTLPATCPFPDQAKHRIERATRMEAGLIVDODDLDE  
EHQGLAPAEFLARLXARMPDGEDLVVTHGDACLPNFIMVNGRPSFGDLCRGLGVADRY  
QDIALATRDIAEELGEWADRFVLVYIGIAAPDSQRIAFYRLLDLDEFF"  
9746..10430  
/note="from mannopine synthase"  
terminator  
ORIGIN  
Query Match 18.5%; Score 1402; DB 11; Length 10450;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6179 GTTTCGGATCTAGATATACATCAATCCACTTGTCTTGAAGACGTGTTGGAACTGCTTTC 6238  
Db 4259 GTTTCGGATCTAGATATACATCAATCCACTTGTCTTGAAGACGTGTTGGAACTGCTTTC 4200  
QY 6239 TTTTTCACGATGTTCTCTGCTGGGTGGGGTCCATCTTTTGGACCACTGTCGTAGAGGC 6298  
Db 4199 TTTTTCACGATGTTCTCTGCTGGGTGGGGTCCATCTTTTGGACCACTGTCGTAGAGGC 4140  
QY 6299 ATCTTGAACGATAGGCTTTCTTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTT 6358  
Db 4139 ATCTTGAACGATAGGCTTTCTTTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTT 4080  
QY 6359 TTCTACTGTCTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCC 6418  
Db 4079 TTCTACTGTCTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCC 4020  
QY 6419 CGATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTT 6478  
Db 4019 CGATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTT 3960  
QY 6479 GATATCTTGGAGTAGAGAGAGTGTCTGCTCCACCATGTTGGGGATCTAGATATCACA 6538  
Db 3959 GATATCTTGGAGTAGAGAGAGTGTCTGCTCCACCATGTTGGGGATCTAGATATCACA 3900  
QY 6539 TCAATCCACTTGTCTTGAAGAGTGTGGAAGCTCTCTTTTCCACCATGTTCTCTCGT 6598  
Db 3899 TCAATCCACTTGTCTTGAAGAGTGTGGAAGAGTCTCTTTTCCACCATGTTCTCTCGT 3840  
QY 6599 GGGTGGGGTCTCATCTTTGGGACCACTGTCGTAGAGGCATCTTGAACCATAGACCTTTCC 6658  
Db 3839 GGGTGGGGTCTCATCTTTGGGACCACTGTCGTAGAGGCATCTTGAACCATAGACCTTTCC 3780  
QY 6659 TTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTCTTTCTACTGTCTTTCGATGAA 6718  
Db 3779 TTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTCTTTCTACTGTCTTTCGATGAA 3720  
QY 6719 GTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTTACCCCTTTGTTGAAA 6778  
Db 3719 GTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTTACCCCTTTGTTGAAA 3660  
QY 6779 AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTTATGATATTTCTTGGAGTAGACGAG 6838  
Db 3659 AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTTATGATATTTCTTGGAGTAGACGAG 3600  
QY 6839 AGTCTGCTGCCACCATGTTGGGATCTAGATATCACAATCAATCACTTGTCTTGAAGA 6898  
Db 3599 AGTCTGCTGCCACCATGTTGGGATCTAGATATCACAATCAATCACTTGTCTTGAAGA 3540  
QY 6899 CGTGGTTGGAGAGTCTTCTTTTCCAGATGTTCTCTGCTGGGTGGGGTCCATCTTTGGG 6958  
Db 3539 CGTGGTTGGAGAGTCTTCTTTTCCAGATGTTCTCTGCTGGGTGGGGTCCATCTTTGGG 3480  
QY 6959 ACCACTGTGGGTAGAGGCATCTTGAACCATAGCCCTTTCTTTATCGCAATGATGGCATTT 7018  
Db 3479 ACCACTGTGGGTAGAGGCATCTTGAACCATAGCCCTTTCTTTATCGCAATGATGGCATTT 3420



```
QY 7019 GTAGAGCCATCTTCCTTTCTACTGTCTCTTCGATGAAGTACACAGATAGCTGGCAATG 7078
Db 3419 GTAGAAGCCATCTTCCTTTCTACTGTCTCTTCGATGAAGTACACAGATAGCTGGCAATG 3360
QY 7079 GAATCCGAGGAGTTTCCGATATATACCTTTGTTGAAAAGTCTCAATAGCCCTCTCGTGC 7138
Db 3359 GAATCCGAGGAGTTTCCGATATATACCTTTGTTGAAAAGTCTCAATAGCCCTCTCGTGC 3300
QY 7139 TTCTGAGACTGTATCTTTGATATTTCTTGAGTACGAGAGTGTCTGCTCCACCATGTT 7198
Db 3299 TTCTGAGACTGTATCTTTGATATTTCTTGAGTACGAGAGTGTCTGCTCCACCATGTT 3240
QY 7199 GGGGATCTAGATATACATCAATCACTTCTTTGAAGACGTGGTGGACGTCTTCTTT 7258
Db 3239 GGGGATCTAGATATACATCAATCACTTCTTTGAAGACGTGGTGGACGTCTTCTTT 3180
QY 7259 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGACACCTGTCTGCTAGAGGCATC 7318
Db 3179 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGACACCTGTCTGCTAGAGGCATC 3120
QY 7319 TTGAACGATAGCTTTCTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTCTTTTC 7378
Db 3119 TTGAACGATAGCTTTCTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTCTTTTC 3060
QY 7379 TACTGTCTTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 7438
Db 3059 TACTGTCTTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 3000
QY 7439 TATTACCTTTCTGGAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTGAT 7498
Db 2999 TATTACCTTTCTGGAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTGAT 2940
QY 7499 ATTCTTGGAGTAGAGAGAGTGTCTGCTCCACCATGTTGGGGATCCACTAGTCTTAGAG 7558
Db 2939 ATTCTTGGAGTAGAGAGAGTGTCTGCTCCACCATGTTGGGGATCCACTAGTCTTAGAG 2880
QY 7559 CGGCCGCCACCGGGTGGAGCT 7580
Db 2879 CGGCCGCCACCGGGTGGAGCT 2858

RESULT 4
AK221402
LOCUS
DEFINITION
Arabidopsis thaliana gene for zinc finger protein OBP3, complete
cds, clone: RAPL25-48-C17.
AK221402
VERSION
AK221402.1 GI:62320151
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS
Totoki,Y., Seki,M., Ishida,J., Nakajima,M., Enju,A., Kamiya,A.,
Narusaka,M., Shin-i,T., Nakagawa,M., Sakamoto,N., Oishi,K.,
Kohara,Y., Kobayashi,M., Toyoda,A., Sakaki,Y., Sakurai,T., Iida,K.,
Akiyama,K., Satou,M., Toyoda,T., Konagaya,A., Carninci,P.,
Kawai,J., Hayashizaki,Y. and Shinozaki,K.
Large-scale analysis of RIKEN Arabidopsis full-length (RAPL) cDNAs
Unpublished
2 (bases 1 to 1457)
TITLE
JOURNAL
REFERENCE
AUTHORS
Totoki,Y., Seki,M., Ishida,J., Nakajima,M., Enju,A., Kamiya,A.,
Narusaka,M., Shin-i,T., Nakagawa,M., Sakamoto,N., Oishi,K.,
Kohara,Y., Kobayashi,M., Toyoda,A., Sakaki,Y., Sakurai,T., Iida,K.,
Akiyama,K., Satou,M., Toyoda,T., Konagaya,A., Carninci,P.,
Kawai,J., Hayashizaki,Y. and Shinozaki,K.
Direct Submission
Submitted (22-MAR-2005) Motoaki Seki, RIKEN Genomic Sciences
Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
230-0045, Japan (E-mail:mseki@gsc.riken.jp
URL:http://range.gsc.riken.jp/, Tel:81-45-503-9625,

QY 7019 GTAGAGCCATCTTCCTTTCTACTGTCTCTTCGATGAAGTACACAGATAGCTGGCAATG 7078
Db 3419 GTAGAAGCCATCTTCCTTTCTACTGTCTCTTCGATGAAGTACACAGATAGCTGGCAATG 3360
QY 7079 GAATCCGAGGAGTTTCCGATATATACCTTTGTTGAAAAGTCTCAATAGCCCTCTCGTGC 7138
Db 3359 GAATCCGAGGAGTTTCCGATATATACCTTTGTTGAAAAGTCTCAATAGCCCTCTCGTGC 3300
QY 7139 TTCTGAGACTGTATCTTTGATATTTCTTGAGTACGAGAGTGTCTGCTCCACCATGTT 7198
Db 3299 TTCTGAGACTGTATCTTTGATATTTCTTGAGTACGAGAGTGTCTGCTCCACCATGTT 3240
QY 7199 GGGGATCTAGATATACATCAATCACTTCTTTGAAGACGTGGTGGACGTCTTCTTT 7258
Db 3239 GGGGATCTAGATATACATCAATCACTTCTTTGAAGACGTGGTGGACGTCTTCTTT 3180
QY 7259 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGACACCTGTCTGCTAGAGGCATC 7318
Db 3179 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGACACCTGTCTGCTAGAGGCATC 3120
QY 7319 TTGAACGATAGCTTTCTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTCTTTTC 7378
Db 3119 TTGAACGATAGCTTTCTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTCTTTTC 3060
QY 7379 TACTGTCTTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 7438
Db 3059 TACTGTCTTTTCGATGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGA 3000
QY 7439 TATTACCTTTCTGGAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTGAT 7498
Db 2999 TATTACCTTTCTGGAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTGAT 2940
QY 7499 ATTCTTGGAGTAGAGAGTGTCTGCTCCACCATGTTGGGGATCCACTAGTCTTAGAG 7558
Db 2939 ATTCTTGGAGTAGAGAGTGTCTGCTCCACCATGTTGGGGATCCACTAGTCTTAGAG 2880
QY 7559 CGGCCGCCACCGGGTGGAGCT 7580
Db 2879 CGGCCGCCACCGGGTGGAGCT 2858

Query Match 9.3%; Score 708; DB 15; Length 1457;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4105 GCAGCAAGGGACCAACATCAGCTAGATGTGTCAACAACGACCAAGAGCTTAATAA 4164
Db 184 GCGCAGAGGGAACCAACATCAGCTAGATGTGTCAACAACGACCAAGAGCTTAATAA 243
QY 4165 TTACCGGAGCTTCTCATCAACCAAGCTTCTCAGGTTGAGGTTGAGTCAAGCTAGAGT 4224
Db 244 TTACCGGAGCTTCTCATCAACCAAGCTTCTCAGGTTGAGGTTGAGTCAAGCTAGAGT 303
QY 4225 GAATTCATGTTGGAGAGCTGCTCGATCGCAAAAGCTCCCATTCCTGAAGAGCTCTAAA 4284
Db 304 GAATTCATGTTGGAGAGCTGCTCGATCGCAAAAGCTCCCATTCCTGAAGAGCTCTAAA 363
QY 4285 TTGCGCTAGATGCTGACTCAACCAATACTAAGTTCTGTTACTTCAATAACTATAGCCTTAC 4344
Db 364 TTGCGCTAGATGCTGACTCAACCAATACTAAGTTCTGTTACTTCAATAACTATAGCCTTAC 423
QY 4345 TCAACCTCGCCATTTCTGCAAAACATGTCGTTCGCTATTGGACACACGTGGGGTTCCTTGAG 4404
Db 424 TCAACCTCGCCATTTCTGCAAAACATGTCGTTCGCTATTGGACACACGTGGGGTTCCTTGAG 483
QY 4405 GAATGTTCTCTGTTGAGGAGGCTTTAGGAGGACAGAGCAAGCAATCCAGATCGAAATC 4464
Db 484 GAATGTTCTCTGTTGAGGAGGCTTTAGGAGGACAGAGCAAGCAATCCAGATCGAAATC 543
QY 4465 TACGGTCTGGTCTCGACTGATAACTACTAGTACTTCAATCACTTCACTTCTTCGCCCAAG 4524
Db 544 TACGGTCTGGTCTCGACTGATAACTACTAGTACTTCAATCACTTCACTTCTTCGCCCAAG 603
QY 4525 TTACTCAAAACCTTAGCAAGTTTCATAGTACGGTCAAAATCCCGAGTTTAAATCCAACTT 4584
Db 604 TTACTCAAAACCTTAGCAAGTTTCATAGTACGGTCAAAATCCCGAGTTTAAATCCAACTT 663
QY 4585 GCCCATCTTGCTCTCTCTCCAAAGCCTTCGAGATTACAAATCAAGCAACACTGGATTAGA 4644
Db 664 GCCCATCTTGCTCTCTCTCCAAAGCCTTCGAGATTACAAATCAAGCAACACTGGATTAGA 723
QY 4645 TTTTGTGGAACTCAAAATAGCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 4704
```

Fax:81-45-503-9586)  
An Arabidopsis full-length cDNA library was constructed essentially  
as reported previously (Seki et al. (1998) Plant J. 15:707-720;  
Seki et al. (2002) Science 296:141-145).  
This clone is in a modified pBluescript vector.  
Please visit our web site (<http://range.gsc.riken.jp/>) for further  
details.

Location/Qualifiers  
1. .1457  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="RAPL25-48-C17"  
/ecotype="Columbia"  
/note="common name: thale cress"  
1. .1457  
/gene="At3g53370"  
134. .1105  
/gene="At3g53370"  
/codon\_start=1  
/product="zinc finger protein OBP3"  
/protein\_id="BAD94355.1"  
/db\_xref="GI:62320152"  
/translation="MVFSSLPVNOFDSOMQQOQHOLECVTTDDQNNPNYLRLQSSP  
KTSOVAGSSQARVNSMVERARIKVPLEALNCPKDSNTKFCYFNYSITOPRHF  
CKTRRYTGGSLRNVPVGGFRNKRKSKSVVSDNTTSTSLTSRPSYN  
PSKHSYQIPEFNSNLPILPLQSLGSDYNSNTGLDFGGTQISNMISGSSGILD  
AMR1PPSOAQOFPPLINTTGLVSSNALYPLLEGVSATQTRNVKASENDQDRDRD  
DGVNLSRNFNLGNIINSRNEEYTSWGNSSWGTFTSTNNSTGHLSP"

ORIGIN  
Query Match 9.3%; Score 708; DB 15; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
Db 724 TTTTGTGGAGCTCAAAATAGCAACATGATAGTGTATGAGTTCATAGTGGGATCTT 783
QY 4705 GGATGATGGAGAAATACCTCCATCACAACAAGCTCAGCAATTCCTTTCTTGATCAACAC 4764
Db 784 GGATGATGGAGAAATACCTCCATCACAACAAGCTCAGCAATTCCTTTCTTGATCAACAC 843
QY 4765 TACCGATGGTGCATCTCTCAACCGGTATATCCATTACTAGAGG 4812
Db 844 TACCGATGGTGCATCTCTCAACCGGTATATCCATTACTAGAGG 891

RESULT 5
AF155818
LOCUS AF155818 1274 bp mRNA linear PLN 23-DEC-1999
DEFINITION Arabidopsis thaliana zinc finger protein OBP3 mRNA, complete cds.
ACCESSION AF155818
VERSION AF155818.1 GI:5059397
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1274)
Kang,H.G. and Singh,K.B.
Characterization of Arabidopsis Dof Transcription Factors, a Novel
Zinc Finger Protein Family in Plants
Unpublished
REFERENCE 2 (bases 1 to 1274)
Kang,H.G. and Singh,K.B.
Direct Submission
TITLE Zinc Finger Protein Family in Plants
JOURNAL Submitted (02-JUN-1999) MCDB, University of California, Los
Angeles, 405 Hilgard Ave., Los Angeles, CA 90095, USA
FEATURES
Source
1..1274
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/ecotype="Columbia"
138..932
/notes="transcriptional activator"
/codon_start=1
/product="zinc finger protein OBP3"
/protein_id="AAD38988.1"
/db_xref="GI:5059398"
/translation="MVVERARIKVLPEAALNCPKDSNTTKFCYFNYSLSLQPRHFC
KTCRRYWTGGSLRNVPGGPRNKRKSKSTVVSTDTNTSTSLTSRPSYSNP
SKFHSYGIQPFNSNLPILPQSLGDIYNSNLTGLDFTGQISNLMISGSGGILDA
WRIPPSQQQPFLLNTTGLVQSSNALYPLLEGVSVATQTRNVKAEENDQDRGRDGD
GVNNLSRNLFGNININSGRNEEYTSWGCNSWTGFTSNNSTGHLSP"

ORIGIN
Query Match 9.3%; Score 703; DB 15; Length 1274;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 703; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4110 AAGGGAAACCAACATCAGCTAGATGTGTCACTACGACAGAACCCCTAATATTACTTAC 4169
Db 16 AAGGGAAACCAACATCAGCTAGATGTGTCACTACGACAGAACCCCTAATATTACTTAC 75
QY 4170 GGCAGCTCTCATCACACCGACTTCTCAGGTTGCGAGTTCGAGTCAAGCTAGAGTCAAT 4229
Db 76 GGCAGCTCTCATCACACCGACTTCTCAGGTTGCGAGTTCGAGTCAAGCTAGAGTCAAT 135
QY 4230 CAATGTGGAAAGCTGCTCGATCGCAAAAGTCCATTGCTGAGCAGCTCTAAATGGC 4289
Db 136 CAATGTGGAAAGCTGCTCGATCGCAAAAGTCCATTGCTGAGCAGCTCTAAATGGC 195
QY 4290 CTAGATGTGACTCAACCAATCTAGTTCCTGTTACTTCAATTAATTAATTAATTAATTA 4349
Db 196 CTAGATGTGACTCAACCAATCTAGTTCCTGTTACTTCAATTAATTAATTAATTAATTA 255
```

```
QY 4350 CTCGCATTCTCGAAAAACATGTCGTCGCTATTGGACACGTGGCGGTTCTCTCAGGAATG 4409
Db 256 CTCGCCATTCTCGAAAAACATGTCGTCGCTATTGGACACGTGGCGGTTCTCTCAGGAATG 315
QY 4410 TTCTGTTGGAGAGGCTTTTAGGAGGAACAAGAGAAAGCAAAATCCAGATCGAAATCTACGG 4469
Db 316 TTCTGTTGGAGAGGCTTTTAGGAGGAACAAGAGAAAGCAAAATCCAGATCGAAATCTACGG 375
QY 4470 TGTGCTCTCGACTGATATATCTACTAGTACTTTCATCTACTTCTCGCCCAAGTTACT 4529
Db 376 TGTGCTCTCGACTGATATATCTACTAGTACTTTCATCTACTTCTCGCCCAAGTTACT 435
QY 4530 CAAACCTAGCAAGTTTCATAGCTACGTCAAATCCCGAGTTTAATTCCAACTTGCCTCA 4589
Db 436 CAAACCTAGCAAGTTTCATAGCTACGTCAAATCCCGAGTTTAATTCCAACTTGCCTCA 495
QY 4590 TCTTGCTCTCTCTCCAAAGCCTTGGAGATTACAATTCAAAGCAACACTGGATTAGATTTG 4649
Db 496 TCTTGCTCTCTCTCCAAAGCCTTGGAGATTACAATTCAAAGCAACACTGGATTAGATTTG 555
QY 4650 GTGGAACTCAATATAGCAACATGATAGTGGTATGATGTTCTAGTGGTGGATCTTGGATG 4709
Db 556 GTGGAACTCAATATAGCAACATGATAGTGGTATGATGTTCTAGTGGTGGATCTTGGATG 615
QY 4710 CATGAGATACCTCCATCACAACAGCTCAGCAATTCCTCTTTTGTATCAACACTACCG 4769
Db 616 CATGAGATACCTCCATCACAACAGCTCAGCAATTCCTCTTTTGTATCAACACTACCG 675
QY 4770 GATTGTGCAATCTTCACACGGGTTATATCCATTACTAGAAAG 4812
Db 676 GATTGTGCAATCTTCACACGGGTTATATCCATTACTAGAAAG 718

RESULT 6
AX508868
LOCUS AX508868 2000 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 3563 from Patent WO216655.
ACCESSION AX508868
VERSION AX508868.1 GI:23390105
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
Harper,J.F., Krebs,J., Wang,X. and Zhu,T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 3563 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
Location/Qualifiers
1..2000
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 4.9%; Score 368; DB 6; Length 2000;
Best Local Similarity 100.0%; Pred. No. 3e+155;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5763 GAATTATGCACTTAATTCAGACTAATCCCCAAATTTTATGATTTTTCG 5822
Db 1 GAATTATGCACTTAATTCAGACTAATCCCCAAATTTTATGATTTTTCG 60
QY 5823 ATTTAATATGTTCACAAATCATATATGCGCAACTAACTTAATTGAAAGACATGAATG 5882
Db 61 ATTTAATATGTTCACAAATCATATATGCGCAACTAACTTAATTGAAAGACATGAATG 120
QY 5883 ACTGAAACCATGCATAATCTCTCAAGTCTCAACCTATGAAGATCATGTAACCAATAGAC 5942
```

Db	121	ACTGAACCATGATATCTCTCAAGTCTCAACCTATGAAGAATCATGTAAACATAGAC	180
Qy	5943	TATCATCATGATTAGTAAATGATGATCTATATATGTTCTTTGACATAGATATGTCAT	6002
Db	181	TATCATCATGATTAGTAAATGATGATCTATATATGTTCTTTGACATAGATATGTCAT	240
Qy	6003	TTATCTGGATATAAAGATGGCGTTTAACTACTTTGCAATTTTGTATATCTTTCTTC	6062
Db	241	TTATCTGGATATAAAGATGGCGTTTAACTACTTTGCAATTTTGTATATCTTTCTTC	300
Qy	6063	TAATACATATGATCAATACACTTTTGTTTTAAAGAAATTAATAAATTTATTCAAAAT	6122
Db	301	TAATACATATGATCAATACACTTTTGTTTTAAAGAAATTAATAAATTTATTCAAAAT	360
Qy	6123	CGATCACA 6130	
Db	361	CGATCACA 368	
RESULT 7			
AB017485/c			
LOCUS			
DEFINITION	AB017485	2939 bp	DNA linear PLN 02-OCT-1999
ACCESSION	AB017485		Nicotiana tabacum gene.
VERSION	AB017485.1	GI:6006285	
KEYWORDS			
SOURCE			Nicotiana tabacum (common tobacco)
ORGANISM			Nicotiana tabacum
REFERENCE			
AUTHORS			Yamada, T., Tanaka, Y. and Kasumi, T.
TITLE			Putative negative regulator for plant apoptosis
JOURNAL			Published Only in DataBase (1999)
REFERENCE			2 (bases 1 to 2939)
AUTHORS			Yamada, T., Tanaka, Y. and Kasumi, T.
TITLE			Direct Submission
JOURNAL			Submitted (03-SEP-1998) Tetsuji Yamada, Okayama University, Faculty of Agriculture, Tsushima Naka 1-1, Okayama, Okayama 700-8530, Japan (E-mail:tdry@cc.okayama-u.ac.jp, Tel:81-86-251-8307, Fax:81-86-251-8307)
FEATURES			
source			1..2939
			/organism="Nicotiana tabacum"
			/mol_type="genomic DNA"
			/cultivar="Bright Yellow"
			/db_xref="taxon:4097"
			/clone="tentatively names as S-5"
			/cell_type="suspension cultures"
ORIGIN			
Query Match			4.8%; Score 362; DB 15; Length 2939;
Best Local Similarity			100.0%; Pred. No. 2.1e-162;
Matches 362; Conservative			0; Mismatches 0; Indels 0; Gaps 0;
Qy	6263	TGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGATCTTGAACGATAGCCTTCCCTT	6322
Db	364	TGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGATCTTGAACGATAGCCTTCCCTT	305
Qy	6323	ATCGCAATGATGCAATTTGGAAGCCATCTTCTTTCTACCTGCTTTCGATGAAGTG	6382
Db	304	ATCGCAATGATGCAATTTGGAAGCCATCTTCTTTCTACCTGCTTTCGATGAAGTG	245
Qy	6383	ACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATATATACCTTTGTTGAAAAGT	6442
Db	244	ACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATATATACCTTTGTTGAAAAGT	185
Qy	6443	CTCAATAGCCCTCTGGTCTTCTGAGACTGTAATCTTTGATATCTTGGAGTAGACGAGT	6502
Db	184	CTCAATAGCCCTCTGGTCTTCTGAGACTGTAATCTTTGATATCTTGGAGTAGACGAGT	125
Qy	6503	GTGCTGCTCCACCATGTTGGGGATCTAGATATACATCAATCCACTTGTCTTGAACGCT	6562
Db	124	GTGCTGCTCCACCATGTTGGGATCTAGATATACATCAATCCACTTGTCTTGAACGCT	65
Qy	6563	GTTTGGACGCTCTCTTTTCCACCATGTTCTCGTGGTGGGTCCATCTTTTGGGACC	6622
Db	64	GTTTGGACGCTCTCTTTTCCACCATGTTCTCGTGGTGGGTCCATCTTTTGGGACC	5
Qy	6623	AC 6624	
Db	4	AC 3	
RESULT 8			
AJ616346/c			
LOCUS			
DEFINITION	AJ616346	1431 bp	DNA linear PLN 01-JUL-2004
ACCESSION	AJ616346		Nicotiana tabacum chloroplast kr1 DNA integrant.
VERSION	AJ616346.1	GI:45724967	
KEYWORDS			
SOURCE			Nicotiana tabacum (common tobacco)
ORGANISM			Nicotiana tabacum
REFERENCE			
AUTHORS			1
TITLE			Huang, C. Y., Avliffe, M. A. and Timmis, J. N.
JOURNAL			Simple and complex nuclear loci created by newly transferred chloroplast DNA in tobacco
PUBMED			15210980
REFERENCE			2 (bases 1 to 1431)
AUTHORS			Huang, C. Y.
TITLE			Direct Submission
JOURNAL			Submitted (05-DEC-2003) Huang C. Y., School of Molecular & Biomedical Science, The University of Adelaide, The University of Adelaide, South Australia, 5005 Adelaide, AUSTRALIA
FEATURES			
Location/Qualifiers			1..1431
			/organism="Nicotiana tabacum"
			/mol_type="genomic DNA"
			/cultivar="Petit Havana (N, N)"
			/db_xref="taxon:4097"
			/note="kr1 integrant"
misc_feature			1..385
			/note="neoTSL2 sequence"
misc_feature			386..451
			/note="chloroplast DNA of pPRV111A"
misc_feature			451..452
			/note="recombination junction (J1-2)"
misc_feature			452..779
			/note="aada sequence of pPRV111A"
misc_feature			729..734
			/note="EcoRI site"
misc_feature			780..1431
			/note="neoTSL2 sequence"
ORIGIN			
Query Match			4.4%; Score 332; DB 15; Length 1431;
Best Local Similarity			100.0%; Pred. No. 6.4e-148;
Matches 332; Conservative			0; Mismatches 0; Indels 0; Gaps 0;
Qy	6190	AGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGACCTCTCTTTTCCACGA	6249
Db	1131	AGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGACCTCTCTTTTCCACGA	1072
Qy	6250	TGTTCTCTGGGTGGGGTCCATCTTTGGGACCACTGTGGAGGATCTTGAACGA	6309
Db	1071	TGTTCTCTGGGTGGGGTCCATCTTTGGGACCACTGTGGAGGATCTTGAACGA	1012
Qy	6310	TAGCCTTTCTTTATCGCAATGATGCAATTTAGAACCATCTTCTTTCTACTGTCC	6369
Db	1011	TAGCCTTTCTTTATCGCAATGATGCAATTTAGAACCATCTTCTTTCTACTGTCC	952

```
QY 6370 TTTCGATGAAGTCACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429
Db 951 TTTCGATGAAGTCACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 892
QY 6430 TTTCGATGAAGTCACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6489
Db 891 TTTCGATGAAGTCACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 832
QY 6490 AGTAGACGAGAGTGTCTGCTCTCCACCATGTTG 6521
Db 831 AGTAGACGAGAGTGTCTGCTCTCCACCATGTTG 800

RESULT 9
VF0551314/c
LOCUS VF0551314 4462 bp DNA circular SYN 27-MAR-2003
DEFINITION Transfection vector pBTdest.
ACCESSION AJ551314
VERSION AJ551314.1 GI:29335742
KEYWORDS amp gene; beta lactamase; cat gene; ccdB gene; chloramphenicol
acetyl transferase; control of cell death B protein.
SOURCE Transfection vector pBTdest
ORGANISM Transfection vector pBTdest
other sequences; artificial sequences; vectors.
REFERENCE 1
AUTHORS Jakob,M.J., Heim,M.A. and Weisshaar,B.
TITLE Use of a gateway compatible vector for transient plant transfection
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4462)
AUTHORS Jakob,M.J.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-2003) Jakob M.J., Salamini, MPI for Plant
Breeding Research, Carl-von-Linne Weg 10, 50829 Koeln, GERMANY
FEATURES
source
1..4462
/organism="Transfection vector pBTdest"
/mol_type="other DNA"
/db_xref="taxon:225975"
31..443
/note="35S"
421..424
/note="35S"
456..580
/note="attR1"
689..1348
/gene="cat"
689..1348
/gene="cat"
/codon_start=1
/product="chloramphenicol acetyl transferase"
/protein_id="CAD83080.1"
/db_xref="GI:29335743"
/translation="MEKKTITGVTVDISOWHRKEHFAFQSVACQYNTQVOLDITAF
LXTVKXKHKFPAPTHILARLMAHPRMAKDGELVIMDSVHPCYTVHEOTETFE
SLWSEYHDDPQFQLHITSQDVACYGENLAYFPKGFIEFMFVSANPWFSTFSLNV
ANMDDFFAPVTMGKYTTGDKVLMPLAIQVHHAHVCDGPHVGRMLNELQQYCYDEWQGG
A"
1690..1995
/gene="ccdB"
1690..1995
/gene="ccdB"
/codon_start=1
/product="control of cell death B protein"
/protein_id="CAD83081.1"
/db_xref="GI:29335744"
/translation="MQFKVYTYKRSRYLFDVQSDIIDTPGRRWIPLASRLISD
KVSRELYPVVHHIGDSRWMTMDMASVPVSVIGEEVADLSHRENDIKNAINLFWGI"
2036..2160
/note="attR2"
2168..2463
/gene="nost"
2168..2463
/gene="nost"

promoter
TATA_signal
misc_recomb
gene
CDS
gene
CDS
gene
misc_recomb
gene
terminator
```

```
gene 2606..3466
CDS /gene="amp"
2606..3466
/gene="amp"
/codon_start=1
/product="beta lactamase"
/protein_id="CAD83082.1"
/db_xref="GI:29335745"
/translation="MSIQHFRVALIPFFAAFLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSKILESFRPEERFPMWSFKVLLCGAVLSRIDAGQEQLGRRIHYSNDLVE
YSPVTEKLTIDGMTVRELCSAITKMSDNTAANLLLTGGPKELTAFLNMGDGHVTRL
DRWEPELNEATPNDERDTTPMVAMATTIKLITGELLTLASRQQLIDMWEADKVAGPL
LRSLAPAGWFIADKSGAGERSGRIIAALGPDGPSRIWVIYTTGSAQTMDERNQIA
EIGASLIKHW"

ORIGIN
Query Match 4.4%; Score 332; DB 11; Length 4462;
Best Local Similarity 100.0%; Pred. No. 5.2e-148;
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6190 AGATATCACATCAATCCACTTCCTTTGAAGACGTGGTTGGAACGCTCTTTTTCACCA 6249
Db 358 AGATATCACATCAATCCACTTCCTTTGAAGACGTGGTTGGAACGCTCTTTTTCACCA 299
QY 6250 TGTTCCTCGTGGGTGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACCA 6309
Db 298 TGTTCCTCGTGGGTGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACCA 239
QY 6310 TAGCCTTTCTTTATCGCAATGATGCGCATTTGTAGAAGCCATCTTCTTTTCTACTGTCC 6369
Db 238 TAGCCTTTCTTTATCGCAATGATGCGCATTTGTAGAAGCCATCTTCTTTTCTACTGTCC 179
QY 6370 TTTCGATGAAGTCACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429
Db 178 TTTCGATGAAGTCACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 119
QY 6430 TTTCGATGAAGTCACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6489
Db 118 TTTCGATGAAGTCACAGATAGCTGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 59
QY 6490 AGTAGACGAGAGTGTCTGCTCTCCACCATGTTG 6521
Db 58 AGTAGACGAGAGTGTCTGCTCTCCACCATGTTG 27

RESULT 10
AF406991/c
LOCUS AF406991 9663 bp DNA linear SYN 13-NOV-2003
DEFINITION Tobacco rattle virus RNA2-based VIGS vector pTRV2, complete
sequence.
ACCESSION AF406991
VERSION AF406991.1 GI:21389151
KEYWORDS Tobacco rattle virus RNA2-based VIGS vector pTRV2 (pYL156)
Tobacco rattle virus RNA2-based VIGS vector pTRV2
SOURCE Other sequences; artificial sequences; vectors.
ORGANISM 1 (bases 1 to 9663)
REFERENCE 1 (bases 1 to 9663)
AUTHORS Liu,Y., Schiff,M., Marathe,R. and Dinesh-Kumar,S.P.
TITLE Tobacco Rar1, EDS1 and NPR1/NIM1 like genes are required for
N-mediated resistance to tobacco mosaic virus
JOURNAL Plant J. 30 (4), 415-429 (2002)
PUBMED 12028572
REFERENCE 2 (bases 1 to 9663)
AUTHORS Liu,Y.
TITLE Direct Submission
JOURNAL Submitted (08-AUG-2001) MCDB, Yale University, 165 Prospect Street,
New Haven, CT 06520, USA
FEATURES
Location/Qualifiers
1..9663
/organism="Tobacco rattle virus RNA2-based VIGS vector
pTRV2"
/mol_type="other DNA"
/db_xref="taxon:188057"
source
```

```
/note="Tobacco rattle virus RNA2-based T-DNA vector
synonym: pYL156"
1..1639
/note="TRV strain ppk20 RNA2 5'-sequence"
1640..1708
/note="multiple cloning site region"
1709..2103
/note="TRV strain ppk20 RNA2 3'-sequence"
2104..2154
/note="synthetic sequence"
2155..2894
/note="pCambia 0390 sequence"
8895..9663
/note="CamV 35S promoter from pCASS2"

ORIGIN
Query Match      4.4%; Score 332; DB 11; Length 9663;
Best Local Similarity 100.0%; Pred. No. 4.5e-148;
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6190 AGATATCATCAATCCACTTCTTTGAAGACGTGGTTGGAACGTCCTTTTTCACGA 6249
Db 9577 AGATATCATCAATCCACTTCTTTGAAGACGTGGTTGGAACGTCCTTTTTCACGA 9518
QY 6250 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
Db 9517 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 9458
QY 6310 TAGCCCTTCTTTATCGCAATGATGCACTTTGTAGAAGCCATCTTCTTTTCTACTGTCC 6369
Db 9457 TAGCCCTTCTTTATCGCAATGATGCACTTTGTAGAAGCCATCTTCTTTTCTACTGTCC 9398
QY 6370 TTTTCGATGAAGTCACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCC 6429
Db 9397 TTTTCGATGAAGTCACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCC 9338
QY 6430 TTTGTTGAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTGATATTTCTGG 6489
Db 9337 TTTGTTGAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTGATATTTCTGG 9278
QY 6490 AGTAGACGAGAGTGTCTGCTCCACCATTGTTG 6521
Db 9277 AGTAGACGAGAGTGTCTGCTCCACCATTGTTG 9246

RESULT 11
AY788908/c      12537 bp      DNA      circular SYN 02-MAR-2005
LOCUS      N-terminal TAPA T-DNA vector pN-TAPA, complete sequence.
DEFINITION      AY788908
ACCESSION      AY788908
VERSION      AY788908.1 GI:55824365
KEYWORDS      .
SOURCE      N-terminal TAPA T-DNA vector pN-TAPA
ORGANISM      N-terminal TAPA T-DNA vector pN-TAPA
other sequences; artificial sequences; vectors.
REFERENCE      1 (bases 1 to 12537)
AUTHORS      Rubio,V., Shen,Y., Saijo,Y., Liu,Y., Gusmaroli,G.,
Dinesh-Kumar,S.P. and Deng,X.W.
An alternative tandem affinity purification strategy applied to
Arabidopsis protein complex isolation
Plant J. 41 (5), 767-778 (2005)
JOURNAL      15703063
PUBMED      2 (bases 1 to 12537)
AUTHORS      Rubio,V. and Deng,X.W.
Direct Submission
TITLE      Submitted (21-OCT-2004) MCDB, Yale University, 165, Prospect St.,
New Haven, CT 06511, USA
JOURNAL      Location/Qualifiers
FEATURES      1..12537
source      /organism="N-terminal TAPA T-DNA vector pN-TAPA"
/mol_type="other DNA"
/db_xref="taxon:300483"
/note="derived from pPZp222"
```

```
promoter
/note="duplicated CamV 35S promoter"
780..844
/note="TMV UI omega sequence"
854..1231
/note="2x protein A IgG binding domain"
1256..1279
/note="protease 3C cleavage site"
1286..1303
/note="6x HIS tag"
1325..1675
/note="9x myc tag"
1704..3577
/note="attR1-CmR-ccdB-attR2 GATEWAY cassette"
3607..3863
/note="NOS terminator"

ORIGIN
Query Match      4.4%; Score 332; DB 11; Length 12537;
Best Local Similarity 100.0%; Pred. No. 4.3e-148;
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6190 AGATATCATCAATCCACTTGTGGAAGACGTGGTTGGAACGTCCTTTTTCACGA 6249
Db 683 AGATATCATCAATCCACTTGTGGAAGACGTGGTTGGAACGTCCTTTTTCACGA 624
QY 6250 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
Db 623 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 564
QY 6310 TAGCCCTTCTTTATCGCAATGATGCACTTTGTAGAAGCCATCTTCTTTTCTACTGTCC 6369
Db 563 TAGCCCTTCTTTATCGCAATGATGCACTTTGTAGAAGCCATCTTCTTTTCTACTGTCC 504
QY 6370 TTTTCGATGAAGTCACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCC 6429
Db 503 TTTTCGATGAAGTCACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTACCC 444
QY 6430 TTTGTTGAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTGATATTTCTGG 6489
Db 443 TTTGTTGAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTGATATTTCTGG 384
QY 6490 AGTAGACGAGAGTGTCTGCTCCACCATTGTTG 6521
Db 383 AGTAGACGAGAGTGTCTGCTCCACCATTGTTG 352

RESULT 12
AY737283/c      12607 bp      DNA      circular SYN 02-MAR-2005
LOCUS      C-terminal TAP T-DNA vector pYL436, complete sequence.
DEFINITION      AY737283
ACCESSION      AY737283
VERSION      AY737283.1 GI:52630860
KEYWORDS      .
SOURCE      C-terminal TAP T-DNA vector pYL436
ORGANISM      C-terminal TAP T-DNA vector pYL436
other sequences; artificial sequences; vectors.
REFERENCE      1 (bases 1 to 12607)
AUTHORS      Rubio,V., Shen,Y., Saijo,Y., Liu,Y., Gusmaroli,G.,
Dinesh-Kumar,S.P. and Deng,X.W.
An alternative tandem affinity purification strategy applied to
Arabidopsis protein complex isolation
Plant J. 41 (5), 767-778 (2005)
JOURNAL      15703063
PUBMED      2 (bases 1 to 12607)
AUTHORS      Liu,Y. and Dinesh-Kumar,S.P.
TITLE      A C-terminal tandem affinity purification (TAP) T-DNA vector pYL436
Unpublished
JOURNAL      3 (bases 1 to 12607)
REFERENCE      Liu,Y. and Dinesh-Kumar,S.P.
Direct Submission
TITLE      Submitted (28-AUG-2004) MCDB, Yale University, New Haven, CT 06520,
USA
JOURNAL      USA
```

FEATURES		Location/Qualifiers	
source	1..12607	/organism="C-terminal TAP T-DNA vector pYL436"	
		/mol_type="other DNA"	
		/db_xref="taxon:293967"	
		/notes="derived from pP222"	
promoter	1..769	/notes="duplicated CamV 35S promoter"	
		/notes="TMV U1 omega sequence"	
misc_feature	780..844		
		/notes="845..2718"	
misc_feature	845..2718		
		/notes="attR1-Cmr-ccdB-attr2 GATEWAY cassette"	
misc_feature	2752..3009		
		/notes="9x myc tag"	
misc_feature	3163..3180		
		/notes="6x HIS tag"	
misc_feature	3181..3204		
		/notes="protease 3C cleavage site"	
misc_feature	3244..3591		
		/notes="2x IGG binding domain"	
misc_feature	3670..3932		
		/notes="Nos terminator"	
ORIGIN			
Query Match 4.4%; Score 332; DB 11; Length 12607;			
Best Local Similarity 100.0%; Pred.No. 4.3e-148;			
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	6190	AGATATCACATCAATCCACTTGTGTTGAAGACGTGTTGGAACGTCTCTTTTTCACGA	6249
Db	683	AGATATCACATCAATCCACTTGTGTTGAAGACGTGTTGGAACGTCTCTTTTTCACGA	624
QY	6250	TGTTCTCTGTTGGTGGGGTCCATCTTTGGACCACTGTCGGTAGAGGCATCTTGAACGA	6309
Db	623	TGTTCTCTGTTGGTGGGGTCCATCTTTGGACCACTGTCGGTAGAGGCATCTTGAACGA	564
QY	6310	TAGCCTTTCCTTATCGCAATGATGTCATTTGAGAGCCATCTTCTTTCTACTGTCC	6369
Db	563	TAGCCTTTCCTTATCGCAATGATGTCATTTGAGAGCCATCTTCTTTCTACTGTCC	504
QY	6370	TTTCGATGAAGTGACAGATAGTCGGCAATGGAATCCGAGGAGTTTCCGATATTACCC	6429
Db	503	TTTCGATGAAGTGACAGATAGTCGGCAATGGAATCCGAGGAGTTTCCGATATTACCC	444
QY	6430	TTTGTTGAAAGTCTCAATAGCCCTCTGTTCTCTGAGACTGTATCTTTGATATTTCTGG	6489
Db	443	TTTGTTGAAAGTCTCAATAGCCCTCTGTTCTCTGAGACTGTATCTTTGATATTTCTGG	384
QY	6490	AGTAGACGAGTGTCTGCTCCACCATGTTG	6521
Db	383	AGTAGACGAGTGTCTGCTCCACCATGTTG	352
RESULT 13			
LOCUS AY234331/c			
DEFINITION Binary vector pLH7500, complete sequence.			
ACCESSION AY234331			
VERSION AY234331.1 GI:29569730			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS Hausmann,L. and Toepfer,R.			
TITLE Development of Plasmid Vectors			
JOURNAL (in) Brauer,D., Roebbelen,G. and Toepfer,R. (Eds.);			
BIOTECHNOLOGY OF CUSTOM-TAILORED RAPE VARIETIES: 155-172;			
GPZ e. V., Von Sieboldstr. 8, Goettingen, Germany (1999)			
REFERENCE			
AUTHORS Hausmann,L. and Toepfer,R.			
TITLE Direct Submission			
JOURNAL Submitted (12-FEB-2003) Institute for Grapevine Breeding			
Geilweilerhof, Siebeldingen 76833, Germany			
Location/Qualifiers			
1..8903			
/organism="Binary vector pLH7500"			
/mol_type="other DNA"			
/db_xref="taxon:226216"			
6..424			
/note="promoter/enhancer sequence from the 35S gene of			
Cauliflower mosaic virus in GenBank Accession Numbers			
X05868 and V00140"			
360..363			
CAAT signal			
TATA signal			
392..398			
misc_feature			
423			
/note="putative transcription start site"			
436..990			
gene			
/gene="bar"			
/note="synonym: pat; bialaphos/phosphinothricin resistance			
gene from Streptomyces hygroscopicus; similar to GenBank			
Accession Number X17220"			
439..990			
CDS			
/gene="bar"			
/note="confers resistance to bialaphos and			
phosphinothricin"			
/codon_start=1			
/transl_table=11			
/product="phosphinothricin acetyl transferase"			
/protein_id="AA08365.1"			
/db_xref="GI:29569731"			
/translation="MGERRPADIRRAEDMPAVCTIVNHVIETSTVNRTEPQEPQ			
EWTDLRLRERYPWIAEVDGCVAGIAGPWKARNAYDWTAESTVYVSRHQRGL			
GSTLYTHLLKSLEAGCFKSVAVIGLPNDPSVRMEALGYAPRGMLRAAGFKHGNWHD			
VGFWQLDFSLPVPFRPVLPTVEI"			
misc_difference 442			
/gene="bar"			
/note="compared to GenBank Accession Number X17220"			
/evidence=experimental			
/replace="a"			
1013..1212			
terminator			
/note="terminator sequence from the 35S gene of			
Cauliflower mosaic virus in GenBank Accession Numbers			
X05868 and V00140"			
misc_difference 1051			
/note="compared to GenBank Accession Number X05868"			
/evidence=experimental			
/replace="a"			
1178..1193			
polya signal			
misc_feature			
1208			
/note="putative transcription stop site"			
1231..1332			
misc_feature			
/note="multiple cloning site from vector pBlueSfi AB in			
GenBank Accession Number AF327874"			
1333..1935			
misc_feature			
/note="right border region from napaline-type Ti plasmid			
C58 of Agrobacterium tumefaciens in GenBank Accession			
Number X07435"			
misc_difference 1358			
/note="compared to GenBank Accession Number X07435"			
/evidence=experimental			
/replace="a"			
1378..1401			
misc_feature			
/note="right flanking T-DNA border"			
repeat_region			
1425..1437			
/note="repeat C"			
/rpt_type=inverted			
misc_difference 1431			
/note="compared to GenBank Accession Number X07435"			
/evidence=experimental			
/replace="a"			
1460..1471			
repeat_region			
/note="repeat C"			
/rpt_type=inverted			
1485..1495			
repeat_region			
/note="repeat B"			



```

SOURCE      Binary vector pLH7000
ORGANISM    Binary vector pLH7000
REFERENCE   Other sequences; artificial sequences; vectors.
AUTHORS     1 (bases 1 to 8911)
TITLE       Hausmann,L. and Toepfer,R.
JOURNAL     Development of Plasmid Vectors
            (in) Brauer,D., Roebelen,G. and Toepfer,R. (Eds.);
            BIOENGINEERING OF CUSTOM-TAILORED RAPE VARIETIES: 155-172;
            GPZ e. V., Von Sieboldstr. 8, Goettingen, Germany (1999)
REFERENCE   2 (bases 1 to 8911)
AUTHORS     Hausmann,L. and Toepfer,R.
TITLE       Direct Submission
JOURNAL     Submitted (12-FEB-2003) Institute for Grapevine Breeding
            Geilweilerhof, Siebeldingen 76833, Germany
FEATURES
  source
    1..8911
      /organism="Binary vector pLH7000"
      /mol_type="other DNA"
      /db_xref="taxon:226215"
    6..424
      /note="promoter/enhancer sequence from the 35S gene of
            Cauliflower mosaic virus in GenBank Accession Numbers
            X05868 and V00140"
    360..363
      CAAT_signal
    392..398
      TATA_signal
    423
      misc_feature
      /note="putative transcription start site"
    436..990
      gene
      /gene="bar"
      /note="synonym: pat; bialaphos/phosphinothricin resistance
            gene from Streptomyces hygroscopicus; similar to GenBank
            Accession Number X17220"
    439..990
      CDS
      /gene="bar"
      /note="confers resistance to bialaphos and
            phosphinothricin"
      /codon_start=1
      /transl_table=11
      /product="phosphinothricin acetyl transferase"
      /protein_id="AAO85360.1"
      /db_xref="GI:29569725"
      /translation="MGPERRPADIRATEADMPAVCTIVNHVYETSTVNFRTPEQPEQ
            EWTDLVLRERYPLWAEVGEVAGIAYAGPWKARNAYDWTAESTVTVSPRKHQETGL
            GSTYLHLKLEAQGFKSVAVIGLPNDPSVRMEALGYAPRGMLRAAGFKGNWHD
            VGFQWLDLSLPVPPRPVLPVTEI"
    442
      misc_difference
      /gene="bar"
      /note="compared to GenBank Accession Number X17220"
      /evidence=experimental
      /replace="a"
    1013..1212
      terminator
      /note="terminator sequence from the 35S gene of
            Cauliflower mosaic virus in GenBank Accession Numbers
            X05868 and V00140"
    1051
      misc_difference
      /note="compared to GenBank Accession Number X05868"
      /evidence=experimental
      /replace="n"
    1178..1193
      polyA_signal
    1208
      misc_feature
      /note="putative transcription stop site"
    1239..1340
      misc_feature
      /note="multiple cloning site from vector pBlueSfi BA in
            GenBank Accession Number AF327875"
    1341..1943
      misc_feature
      /note="right border region from nopaline-type Ti plasmid
            C58 of Agrobacterium tumefaciens in GenBank Accession
            Number X07435"
    1366
      misc_difference
      /note="compared to GenBank Accession Number X07435"
      /evidence=experimental
      /replace="a"
    1386..1409
      misc_feature

```

---

```

repeat_region
1433..1445
  /note="right flanking T-DNA border"
  /note="repeat C"
  /rpt_type=inverted
misc_difference
1439
  /note="compared to GenBank Accession Number X07435"
  /evidence=experimental
  /replace="n"
repeat_region
1468..1479
  /note="repeat C"
  /rpt_type=inverted
repeat_region
1493..1503
  /note="repeat B"
  /rpt_type=inverted
repeat_region
1520..1530
  /note="repeat B"
  /rpt_type=inverted
misc_difference
1525
  /note="compared to GenBank Accession Number X07435"
  /evidence=experimental
  /replace="n"
misc_difference
1536
  /note="compared to GenBank Accession Number X07435"
  /evidence=experimental
  /replace="a"
misc_feature
  complement(1553..1943)
  /note="3' sequence of the ornithine cyclodeaminase gene
        ocd; non-functional"
misc_difference
1729
  /note="compared to GenBank Accession Number X07435"
  /evidence=experimental
  /replace="t"
rep_origin
1944..3108
  /note="ColEI origin of replication sequence of pBR322;
        similar to GenBank Accession Number J01749; serves as
        origin of replication in Escherichia coli host"
  /evidence=experimental
misc_difference
2201
  /note="compared to GenBank Accession Number J01749"
  /evidence=experimental
  /replace="c"
misc_difference
2740..2741
  /note="compared to GenBank Accession Number J01749"
  /evidence=experimental
  /replace="gta"
misc_feature
2920
  /note="pic/bom site"
rep_origin
3124..6895
  /note="origin of replication and partitioning region from
        plasmid pV81 of Pseudomonas; similar to GenBank Accession
        Number U10487; serves as origin of replication in
        Agrobacterium tumefaciens host"
  /evidence=experimental
-35_signal
3255..3260
-10_signal
3281..3286
RBS
3303..3307
CDS
3317..4003
  /note="resolvase-like protein encoded by parC of plasmid
        pV81; similar to Tn3 resolvase in GenBank Accession Number
        X00613, Tn917 resolvase in GenBank Accession Number
        M11180, and RK2 Para in GenBank Accession Number L27758"
  /codon_start=1
  /transl_table=11
  /product="resolvase"
  /protein_id="AAO85361.1"
  /db_xref="GI:29569726"
  /translation="MNKSAAGLLGYARVSTDDQLTNQRAELHAAGCTKLPSEKITG
        TRRPELARMDLHLPQGVVTVTRDLRLARSTRDLDAERIQAAGALSLAEPWA
        DTTTPAGRMVLTVFAGIAEFERSLIIDTRSGREAARKGVKFGPRPTLTPAQIAHAR
        ELIQEGRTVKEAALLGVHRSTLYRALERSEVTPTEARRRGAFREDALTEADALAA
        AENERQEQEA"
misc_difference
3736
  /note="original nucleotide changed to destroy a Sfil

```



```
restriction site"
/evidence=experimental
/replace="c"
4147..4179
/note="palindromic sequence"
misc_difference 4155
/note="nucleotide changed to conserve palindromic sequence
following SfiI restriction site-destruction"
/evidence=experimental
/replace="g"
misc_difference 4170
/note="original nucleotide changed to destroy two SfiI
restriction sites"
/evidence=experimental
/replace="c"
4237..4242
4258..4263
4291..4298
4302..4931
/note="partitioning protein ParA encoded by parA of
plasmid pVSI; similar to F plasmid SopA in GenBank
Accession Number X04619, RK2 IncC in GenBank Accession
Number L27758, and P1 ParA in GenBank Accession Number
X02954"
/codon_start=1
/transl_table=11
/product="ParA"
/protein_id="AA085362.1"
/db_xref="GI:29569728"
/translation="MKVIAVLNQKSGKTTIATHALQLAGDVLVSDPQSGA
RDMAAVREDQPLTVVIGIDRPTVDVKAIGRDFVVIDGAPOADLVAISAIAADFLV
IPVSPYDINATDELVKQRIEIVTRLOAFVSRRAIKGTIGGVEAEALAGYE
LPILSRITQVSYPCPTAAAGTTVLESPGDAAREVQALAAEIKSLI"
5198..5210
/note="identical to the korB operator binding site of
plasmid RK2 in GenBank Accession Number L27758.1"
5258..6433
/note="partitioning protein ParB encoded by parB of
plasmid pVSI; similar to F plasmid SopB in GenBank
Accession Number X04619, RK2 KorB in GenBank Accession
Number L27758, and P1 ParB in GenBank Accession Number
X02954"
/codon_start=1
/transl_table=11
/product="ParB"
/protein_id="AA085363.1"
/db_xref="GI:29569728"
/translation="MPHWNRGLRKLRLKGLCLPALQWHPQARGIGVSGRKPSPGV
QIGAAUGDDLVEKLKAAQAAQRIEAEARPGESWQAAADRIKESRQPPAAGAFSIR
4.4%; Score 331; DB 11; Length 8911;
Best Local Similarity 100.0%; Pred. No. 1.4e-147;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query Match
6190 AGATATCACATCAATCCACTGCTTTGAAGACGTGGTGGAACTCTTTTTCACCA 6249
336 AGATATCACATCAATCCACTGCTTTGAAGACGTGGTGGAACTCTTTTTCACCA 277
6250 TGTTCTCTGGTGGGGTCCATCTTTGGGACCATCTGCGTAGAGGCATCTTGAACCA 6309
276 TGTTCTCTGGTGGGGTCCATCTTTGGGACCATCTGCGTAGAGGCATCTTGAACCA 217
6310 TAGCCTTTCTTTATCGCAATGATGCGATTTGAGAACCATCTCTCTTTCTACTGTCC 6369
216 TAGCCTTTCTTTATCGCAATGATGCGATTTGAGAACCATCTCTCTTTCTACTGTCC 157
6370 TTTTCGATGAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCC 6429
156 TTTTCGATGAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATTACCC 97
6430 TTTGTTGAAAGTCTCATAGACCTCTGGTCTTCTCGAGCTGTATCTTTGATATTTCTGG 6489
96 TTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTCGAGCTGTATCTTTGATATTTCTGG 37
```

```
QY 6490 AGTAGACGAGAGTGTCTGCTCCACCATGTT 6520
|||||
Db 36 AGTAGACGAGAGTGTCTGCTCCACCATGTT 6
|||||

RESULT 15
AY234327/c 8958 bp DNA circular SYN 10-APR-2003
LOCUS AY234327
DEFINITION Binary vector pLH5000, complete sequence.
ACCESSION AY234327
VERSION AY234327.1 GI:29569706
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 8958
/organism="Binary vector pLH5000"
/mol_type="other DNA"
/db_xref="taxon:226212"
6..424
/note="promoter/enhancer sequence from the 35S gene of
Cauliflower mosaic virus in GenBank Accession Numbers
X05868 and V00140"
360..363
392..398
423
/note="putative transcription start site"
439..1051
/gene="dhfr"
/note="synonym: mtx; methotrexate resistance gene from Mus
musculus; similar to GenBank Accession Number V00734"
446..1009
/gene="dhfr"
/note="confers resistance to methotrexate"
1
/codon_table=11
/product="dihydrofolate reductase"
/protein_id="AA085345.1"
/db_xref="GI:29569707"
/translation="MVRPLNCIVAVSQNMIGKNGDRPWPPLRNEFKYFQMTTSSV
EGKQNLVIMGRKTFISIPKRNPLKDRINIVLSRELKPPRGHFLAKSLDLDALRIE
QPELASKVDWMIWVGSSVYQAMNQPHLRFLVTRIMQEFESDTFFPBDILGKYKLL
PEIPGVLSVEVEQEKIKKPYEYKDD"
1060..1259
/note="terminator sequence from the 35S gene of
Cauliflower mosaic virus in GenBank Accession Numbers
X05868 and V00140"
misc_difference 1098
/note="compared to GenBank Accession Number X05868"
/evidence=experimental
/replace="n"
1225..1240
1255
/note="putative transcription stop site"
1286..1387
/note="multiple cloning site from vector pBlueSfi BA in
GenBank Accession Number AF327875"
1388..1990
/note="right border region from nopaline-type Ti plasmid
C58 of Agrobacterium tumefaciens in GenBank Accession
```

```

Number X07435"
misc_difference 1413 /note="compared to GenBank Accession Number X07435"
/evidence=experimental
/replaces="a"
misc_feature 1433..1456 /note="right flanking T-DNA border"
repeat_region 1480..1492 /note="repeat C"
/rpt_type=inverted
misc_difference 1486 /note="compared to GenBank Accession Number X07435"
/evidence=experimental
/replaces="n"
repeat_region 1515..1526 /note="repeat C"
/rpt_type=inverted
repeat_region 1540..1550 /note="repeat B"
/rpt_type=inverted
repeat_region 1567..1577 /note="repeat B"
/rpt_type=inverted
misc_difference 1572 /note="compared to GenBank Accession Number X07435"
/evidence=experimental
/replaces="n"
misc_difference 1583 /note="compared to GenBank Accession Number X07435"
/evidence=experimental
/replaces="a"
misc_feature complement(1600..1990) /note="3' sequence of the ornithine cyclodeaminase gene
cod; non-functional"
misc_difference 1776 /note="compared to GenBank Accession Number X07435"
/evidence=experimental
/replaces="t"
rep_origin 1991..3155 /note="ColE1 origin of replication sequence of pBR322;
similar to GenBank Accession Number J01749; serves as
origin of replication in Escherichia coli host"
/evidence=experimental
misc_difference 2248 /note="compared to GenBank Accession Number J01749"
/evidence=experimental
/replaces="c"
misc_difference 2787..2788 /note="compared to GenBank Accession Number J01749"
/evidence=experimental
/replaces="gta"
misc_feature 2967 /note="nic/bom site"
rep_origin 3171..6942 /note="origin of replication and partitioning region from
plasmid pVSI1 of Pseudomonas; similar to GenBank Accession
Number U10487; serves as origin of replication in
Agrobacterium tumefaciens host"
/evidence=experimental
-35_signal 3302..3307
-10_signal 3328..3333
RBS 3350..3354
CDS 3364..4050
/note="resolvase-like protein encoded by parR of plasmid
pVSI1; similar to Th3 resolvase in GenBank Accession Number
V00613, Th917 resolvase in GenBank Accession Number
M1180, and RK2 ParA in GenBank Accession Number L27758"
/codon_start=1
/transl_table=11
/product="resolvase"
/protein_id="AA085346.1"
/db_xref="GI:29569708"
/translation="MNKSAAGLLGYARVSTDDQDLTNQRAELHAAGCTKLFSEKITG

TRDRPELARMDLHLRPGDVVTVTRLDRLARSTRDLDIAERIQAEGAGRLSIAEPWA
DTTTPAGRMVLTVPAGIAEFESLIIDRTSRGEAAKARGVKFGEPTTTPAQIAHAR
ELIDQGEIVTKAAALLGVRSTLRALRSSEVFTTEARRRGAFREDALTADALAA
AENEROEQA"
misc_difference 3783 /note="original nucleotide changed to destroy a SfiI
restriction site"
/evidence=experimental
/replaces="c"
stem_loop 4194..4226
/note="palindromic sequence"
misc_difference 4202 /note="nucleotide changed to conserve palindromic sequence
following SfiI restriction site-destruction"
/evidence=experimental
/replaces="g"
misc_difference 4217 /note="original nucleotide changed to destroy two SfiI
restriction sites"
/evidence=experimental
/replaces="c"
-35_signal 4284..4289
-10_signal 4305..4310
RBS 4338..4345
CDS 4349..4978
/note="partitioning protein ParA encoded by parA of
plasmid pVSI1; similar to F plasmid SopA in GenBank
Accession Number X04619, RK2 IncC in GenBank Accession
Number L27758, and P1 ParA in GenBank Accession Number
X02954"
/codon_start=1
/transl_table=11
/product="ParA"
/protein_id="AA085347.1"
/db_xref="GI:29569709"
/translation="MKVIATLKNQKGGSKTTIATLARALQLAGADVLVDSPPQSSA
RDMAAVREDQPLTVVGDIDPPTDRDVKAGREDVFVVDGAPQADLAVSAIKAADFVL
IPVQSPDITWATDLVELVKORIEVTDRLQAAFFVSRAIKTRIGGVEAEALAGYE
LPILSPRITQRVSYPGTAAAGTVLESEPEGAAREVQALAAEIKLI"
misc_feature 5245..5257
/note="identical to the korB operator binding site of
plasmid RK2 in GenBank Accession Number L27758.1"
CDS 5305..6480
/note="partitioning protein ParB encoded by parB of
plasmid pVSI1; similar to F plasmid SopB in GenBank
Accession Number X04619, RK2 KorB in GenBank Accession
Number L27758, and P1 ParB in GenBank Accession Number
X02954"
/codon_start=1
/transl_table=11
/product="ParB"
/protein_id="AA085348.1"
/db_xref="GI:29569710"
/translation="MPVWRNGRLARRKRLGCLPALQWHNPOARGIGVSGRKPSGPV
QIGALGDDLVKELKAAQARQIEAEAPGESWQAAADRIRKESRQPPAACAPSR
KPPKGEQDPFFPMLYDVGTDRSDSIMVAFRLSKRRRAGEVIRYELPDGHVVS
APAGVASVMDYDLVLMVASHLTSNNRYREGKDPGRVFRPHVADLVKFCRRADGG
KOKDLVETCIRLNTTHVAMQTKKAGRLVTVSEGEALISYKIVKSETGPEYIE
IELADVMVREITTEGNPDVLTVPDYLDPDGTGRFLYLRARRAAGKAEARWLFTIY
ERSGSAGEFKFCFTVRKLTGSDNLPYDLKEAGQAGPILVMRYRNLTIEGEASGS"
misc_difference 5411 /note="compared to GenBank Accession Number U10487"

Query Match 4.4%; Score 331; DB 11; Length 8958;
Best Local Similarity 100.0%; Pred. No. 1.4e-147;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6190 AGATATCATCATCAATCCACTTCTTTGAAGACGTGTTGAACGCTCTTTTTCACGA 6249
Db 336 AGATATCATCATCAATCCACTTCTTTGAAGACGTGTTGAACGCTCTTTTTCACGA 277
QY 6250 TGTTCCTCGTGGTGGGTCCATCTTTGGACCACCTGCGTAGAGGATCTTTGAACGA 6309

```

Db 276 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAAACGA 217  
QY 6310 TAGCCCTTTCCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCTTTTCTACTGTCC 6369  
Db 216 TAGCCCTTTCCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCCTTTTCTACTGTCC 157  
QY 6370 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCC 6429  
Db 156 TTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCC 97  
QY 6430 TTTGTGAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTTGATATTCTTGG 6489  
Db 96 TTTGTGAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTTGATATTCTTGG 37  
QY 6490 AGTAGACGAGAGTGTGCTGCTCCACCATTGT 6520  
Db 36 AGTAGACGAGAGTGTGCTGCTCCACCATTGT 6

Search completed: December 31, 2005, 14:03:29  
Job time : 24339 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2005, 06:23:49 ; Search time 841 Seconds  
(without alignments)  
16021.305 Million cell updates/sec

Title: US-10-650-249-1  
Perfect score: 7580  
Sequence: 1 agctctattaattcaagaga.....gcgcacccgcggtgagct 7580

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size: 6

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/1\_COMB.seq:  
2: /cgn2\_6/prodata/1/ina/5\_COMB.seq:  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:  
5: /cgn2\_6/prodata/1/ina/H\_COMB.seq:  
6: /cgn2\_6/prodata/1/ina/PCUS\_COMB.seq:  
7: /cgn2\_6/prodata/1/ina/PP\_COMB.seq:  
8: /cgn2\_6/prodata/1/ina/RE\_COMB.seq:  
9: /cgn2\_6/prodata/1/ina/backfileseq1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	299	3.9	323	3	US-10-012-070A-46
C 2	281	3.7	4839	3	US-09-353-332-1
C 3	222	2.9	470	3	US-09-469-211A-19
C 4	222	2.9	5767	3	US-09-810-861B-3
C 5	222	2.9	14446	3	US-09-810-861B-4
C 6	221	2.9	2728	3	US-08-836-402B-7
C 7	175	2.3	532	3	US-09-042-426-1
C 8	175	2.3	532	3	US-09-291-238-1
C 9	175	2.3	532	3	US-09-330-760-1
C 10	175	2.3	532	3	US-09-328-473-1
C 11	175	2.3	532	3	US-09-330-737-1
C 12	175	2.3	532	3	US-09-329-169-1
C 13	175	2.3	532	3	US-09-330-714A-1
C 14	175	2.3	532	3	US-09-328-826-1
C 15	175	2.3	532	3	US-09-289-170-1
C 16	175	2.3	7378	3	US-09-042-426-9
C 17	175	2.3	7378	3	US-09-291-238-9
C 18	175	2.3	7378	3	US-09-330-760-9
C 19	175	2.3	7378	3	US-09-328-473-9
C 20	175	2.3	7378	3	US-09-330-737-9
C 21	175	2.3	7378	3	US-09-329-169-9
C 22	175	2.3	7378	3	US-09-330-714A-9
C 23	175	2.3	7378	3	US-09-328-826-9
C 24	175	2.3	7378	3	US-09-289-170-9

C 25	98	1.3	12614	3	US-09-577-424-1	Sequence 1, Appli
C 26	97	1.3	332	3	US-09-028-819-11	Sequence 11, Appl
C 27	97	1.3	332	3	US-09-556-163-11	Sequence 11, Appl
C 28	97	1.3	499	3	US-09-028-819-12	Sequence 12, Appl
C 29	97	1.3	499	3	US-09-556-163-12	Sequence 12, Appl
C 30	97	1.3	978	2	US-08-446-486-31	Sequence 31, Appl
C 31	97	1.3	978	2	US-08-463-308-31	Sequence 31, Appl
C 32	97	1.3	979	2	US-08-446-486-30	Sequence 30, Appl
C 33	97	1.3	979	2	US-08-463-308-30	Sequence 30, Appl
C 34	97	1.3	980	9	5254799-30	Patent No. 5254799
C 35	97	1.3	1034	3	US-09-363-970-35	Sequence 35, Appl
C 36	97	1.3	1303	3	US-08-894-440-2	Sequence 2, Appli
C 37	97	1.3	1303	3	US-09-458-093-2	Sequence 2, Appli
C 38	97	1.3	1600	3	US-09-028-819-14	Sequence 14, Appl
C 39	97	1.3	1600	3	US-09-556-163-14	Sequence 14, Appl
C 40	97	1.3	5033	2	US-08-038-768A-1	Sequence 1, Appli
C 41	97	1.3	7811	2	US-08-549-680A-5	Sequence 5, Appli
C 42	96	1.3	831	3	US-08-450-834-5	Sequence 5, Appli
C 43	96	1.3	4544	3	US-09-488-270A-1	Sequence 1, Appli
C 44	96	1.3	4583	3	US-08-810-720-9	Sequence 9, Appli
C 45	93	1.2	439	2	US-08-247-809A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-10-012-070A-46/c  
; Sequence 46, Application US/10012070A  
; Patent No. 6867293  
; GENERAL INFORMATION:  
; APPLICANT: Hawkes, Timothy  
; APPLICANT: Warner, Simon  
; APPLICANT: Andrews, Christopher  
; APPLICANT: Bachoo, Satvinder  
; APPLICANT: Pickerill, Andrew  
; TITLE OF INVENTION: Herbicide Resistant Plants  
; FILE REFERENCE: 50490/UST  
; CURRENT APPLICATION NUMBER: US/10/012,070A  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: PCT/GB00/01573  
; PRIOR FILING DATE: 2000-04-20  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 46  
; LENGTH: 323  
; TYPE: DNA  
; ORGANISM: Cauliflower mosaic virus  
US-10-012-070A-46

Query Match	3.9%	Score 299;	DB 3;	Length 323;
Best Local Similarity	100.0%;	Pred. No. 7.8e-104;		
Matches 299;	Conservative 0;	Mismatches 0;	Gaps 0;	
QY	6198	CATCAATCCACTGCTTTGAACACGCTGGTGGAACTCTCTTTTCCAGATGTTCTTC	6257	
Db	323	CATCAATCCACTGCTTTGAACACGCTGGTGGAACTCTCTTTTCCAGATGTTCTTC	264	
QY	6258	GTGGGTGGGGGCCATCTTTGGGACCACTGCGTAGAGGCATCTTTGAACGATAGCCTTT	6317	
Db	263	GTGGGTGGGGGCCATCTTTGGGACCACTGCGTAGAGGCATCTTTGAACGATAGCCTTT	204	
QY	6318	CCTTTATCGAATGATGGCAATTTAGAACCATCTCTCTTTCTACTCTGCTTTCCGATG	6377	
Db	203	CCTTTATCGAATGATGGCAATTTAGAACCATCTCTCTTTCTACTCTGCTTTCCGATG	144	
QY	6378	AGTGACAGATAGCTGGCAATGGAAATCCGAGAGTTTCCCGATATTACCTTTGTTGA	6437	
Db	143	AGTGACAGATAGCTGGCAATGGAAATCCGAGAGTTTCCCGATATTACCTTTGTTGA	84	
QY	6438	AAAGTCTCAATAGCCCTCTGCTTCTGAGACTGTATCTTTGATATCTTTGGAGTAGAC	6496	
Db	83	AAAGTCTCAATAGCCCTCTGCTTCTGAGACTGTATCTTTGATATCTTTGGAGTAGAC	25	

```
; OTHER INFORMATION: Description of Artificial Sequence: chimaeric
; OTHER INFORMATION: promoter
US-09-469-211A-19

Query Match      2.9%; Score 222; DB 3; Length 470;
Best Local Similarity 100.0%; Pred. No. 1.2e-74;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6300 TCTTGAACGATAGACCTTTCCCTTTATCGCAATGATGGCAATTTGTAGAGCCCATCTTCCCTTT 6359
      |||
Db 235 TCTTGAACGATAGACCTTTCCCTTTATCGCAATGATGGCAATTTGTAGAGCCCATCTTCCCTTT 176

QY 6360 TCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC 6419
      |||
Db 175 TCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC 116

QY 6420 GATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTG 6479
      |||
Db 115 GATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTG 56

QY 6480 ATATTCTTGGAGTAGACGAGAGTGTCGTGCTCCACCATGTTG 6521
      |||
Db 55 ATATTCTTGGAGTAGACGAGAGTGTCGTGCTCCACCATGTTG 14

RESULT 4
US-09-810-861B-3/c
; Sequence 3, Application US/09810861B
; Patent No. 6770799
; GENERAL INFORMATION:
; APPLICANT: Mor, Tsafir S.
; APPLICANT: Soreq, Hermona
; APPLICANT: Arntzen, Charles J.
; APPLICANT: Mabon, Hugh S.
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN
; TITLE OF INVENTION: TRANSGENIC PLANTS
; FILE REFERENCE: BTI-45
; CURRENT APPLICATION NUMBER: US/09/810,861B
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/190,440
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3
; LENGTH: 5767
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid vector
; OTHER INFORMATION: pTM034.
US-09-810-861B-3

Query Match      2.9%; Score 222; DB 3; Length 5767;
Best Local Similarity 100.0%; Pred. No. 7.9e-75;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6300 TCTTGAACGATAGACCTTTCCCTTTATCGCAATGATGGCAATTTGTAGAGCCCATCTTCCCTTT 6359
      |||
Db 241 TCTTGAACGATAGACCTTTCCCTTTATCGCAATGATGGCAATTTGTAGAGCCCATCTTCCCTTT 182

QY 6360 TCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC 6419
      |||
Db 181 TCTACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC 122

QY 6420 GATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTG 6479
      |||
Db 121 GATATTACCCCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTG 62

QY 6480 ATATTCTTGGAGTAGACGAGAGTGTCGTGCTCCACCATGTTG 6521
      |||
Db 61 ATATTCTTGGAGTAGACGAGAGTGTCGTGCTCCACCATGTTG 20
```

## RESULT 5

US-09-810-861B-4/c  
; Sequence 4, Application US/09810861B  
; Patent No. 6770799  
; GENERAL INFORMATION:  
; APPLICANT: Mot. Teafir S.  
; APPLICANT: Soreq, Hermona  
; APPLICANT: Arentzen, Charles J.  
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN  
; TITLE OF INVENTION: TRANSGENIC PLANTS  
; FILE REFERENCE: BTI-45  
; CURRENT APPLICATION NUMBER: US/09/810,861B  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/190,440  
; PRIOR FILING DATE: 2000-03-17  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 4  
; LENGTH: 14446  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (11862)..(12157)  
; OTHER INFORMATION: Description of Artificial Sequence: plasmid vector  
; OTHER INFORMATION: pTM036. Identity of sequence residues 11862-12157 unknown.  
US-09-810-861B-4

Query Match 2.9%; Score 222; DB 3; Length 14446;

Best Local Similarity 100.0%; Pred. No. 6.7e-75; Indels 0; Gaps 0;  
Matches 222; Conservative 0; Mismatches 0;  
QY 6300 TCTTGAACGATAGCCCTTCCTTTATCGCAATGATGCAATTTGTAGAAGCCATCTTCCTTT 6359  
Db 13991 TCTTGAACGATAGCCCTTCCTTTATCGCAATGATGCAATTTGTAGAAGCCATCTTCCTTT 13932  
QY 6360 TCTACTGCTCTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC 6419  
Db 13931 TCTACTGCTCTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC 13872  
QY 6420 GATATTACCTTTGTTGAAAGCTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTG 6479  
Db 13871 GATATTACCTTTGTTGAAAGCTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTG 13812  
QY 6480 ATATTCTTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTTG 6521  
Db 13811 ATATTCTTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTTG 13770

## RESULT 6

US-08-836-402B-7/c  
; Sequence 7, Application US/08836402B  
; Patent No. 6063988  
; GENERAL INFORMATION:  
; APPLICANT: Rudiger Hain, Regina Fischer  
; TITLE OF INVENTION: DNA SEQUENCE AND ITS USE  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE  
; STREET: 660 White Plains Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10591-5144  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage  
; COMPUTER: Gateway 2000  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Microsoft Windows 98  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/836,402B  
; FILING DATE: 02-SEP-1997

CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 44 40 200.7 (Germany)  
; FILING DATE: 10-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kurt G. Briscoe  
; REGISTRATION NUMBER: 33,141  
; REFERENCE/DOCKET NUMBER: Bayer 9840-KGB  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (914) 332-1700  
; TELEFAX: (914) 332-1844  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2728 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
US-08-836-402B-7  
Query Match 2.9%; Score 221; DB 3; Length 2728;  
Best Local Similarity 100.0%; Pred. No. 2.1e-74; Indels 0; Gaps 0;  
Matches 221; Conservative 0; Mismatches 0;  
QY 6300 TCTTGAACGATAGCCCTTCCTTTATCGCAATGATGCAATTTGTAGAAGCCATCTTCCTTT 6359  
Db 516 TCTTGAACGATAGCCCTTCCTTTATCGCAATGATGCAATTTGTAGAAGCCATCTTCCTTT 457  
QY 6360 TCTACTGCTCTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC 6419  
Db 456 TCTACTGCTCTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCC 397  
QY 6420 GATATTACCTTTGTTGAAAGCTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTG 6479  
Db 396 GATATTACCTTTGTTGAAAGCTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTG 337  
QY 6480 ATATTCTTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTT 6520  
Db 336 ATATTCTTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTT 296  
RESULT 7  
US-09-042-426-1/c  
; Sequence 1, Application US/09042426  
; Patent No. 6114608  
; GENERAL INFORMATION:  
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi  
; TITLE OF INVENTION: DNA Construct Containing Bacillus  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6114608artis Corporation  
; STREET: 564 Morris Avenue  
; CITY: Summit  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07901  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/042,426  
; FILING DATE: March 13, 1998  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hoxie, Thomas  
; REGISTRATION NUMBER: 32,993  
; REFERENCE/DOCKET NUMBER: 135/1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919) 541-8614

```
TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: 35S Promoter
;
US-09-042-426-1
Query Match 2.3%; Score 175; DB 3; Length 532;
Best Local Similarity 99.3%; Pred. No. 7.9e-57;
Matches 275; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTCTTGAAGACGCTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 289 AGATATCACATCAATCCACTTCTTGAAGACGCTGGTTGGAACGTCCTCTTTTCCACGA 230
QY 6250 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
Db 229 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 170
QY 6310 TAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCC 6369
Db 169 TAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCC 110
QY 6370 TTTGATGAAGTAGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATACCC 6429
Db 109 TTTGATGAAGTAGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATACCC 50
QY 6430 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTTCTGA 6466
Db 49 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTTCTGA 13

RESULT 8
US-09-291-238-1/c
; Sequence 1, Application US/09291238
; Patent No. 6222104
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 6222104artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/291,238
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/042,426
; FILING DATE: March 13, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoxie, Thomas
; REGISTRATION NUMBER: 32,993
; REFERENCE/DOCKET NUMBER: 135/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8614
; TELEFAX: (919) 541-8689

TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: 35S Promoter
;
US-09-291-238-1
Query Match 2.3%; Score 175; DB 3; Length 532;
Best Local Similarity 99.3%; Pred. No. 7.9e-57;
Matches 275; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTCTTGAAGACGCTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 289 AGATATCACATCAATCCACTTCTTGAAGACGCTGGTTGGAACGTCCTCTTTTCCACGA 230
QY 6250 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
Db 229 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 170
QY 6310 TAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCC 6369
Db 169 TAGCCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTCTACTGTCC 110
QY 6370 TTTGATGAAGTAGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATACCC 6429
Db 109 TTTGATGAAGTAGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATACCC 50
QY 6430 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTTCTGA 6466
Db 49 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTTCTGA 13

RESULT 9
US-09-330-760-1/c
; Sequence 1, Application US/09330760
; Patent No. 6229075
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 6229075artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/330,760
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/042,426
; FILING DATE: March 13, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoxie, Thomas
; REGISTRATION NUMBER: 32,993
; REFERENCE/DOCKET NUMBER: 135/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8614
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
```



```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: 35S Promoter
;
US-09-330-760-1
Query Match          2.3%; Score 175; DB 3; Length 532;
Best Local Similarity 99.3%; Pred. No. 7.9e-57;
Matches 275; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTGGAAGACGTGGTGGAAACGTCCTTTTTCACGGA 6249
Db 289 AGATATCACATCAATCCACTTGTGGAAGACGTGGTGGAAACGTCCTTTTTCACGGA 230
QY 6250 TGTTCCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTGAACGA 6309
Db 229 TGTTCCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTGAACGA 170
QY 6310 TAGCCTTTCCCTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTCTACTGTCC 6369
Db 169 TAGCCTTTCCCTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTCTACTGTCC 110
QY 6370 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429
Db 109 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 50
QY 6430 TTTGTGTAAGTCTCAATAGCCCTCTGCTCTTCTGA 6466
Db 49 TTTGTGTAAGTCTCAATAGCCCTCTGCTCTTCTGA 13
```

```

RESULT 10
US-09-328-473-1/c
; Sequence 1, Application US/09328473
; Patent No. 6232533
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 623253artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,473
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/042,426
; FILING DATE: March 13, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoxie, Thomas
; REGISTRATION NUMBER: 32,993
; REFERENCE/DOCKET NUMBER: 135/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8614
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
```

```

; LENGTH: 532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: 35S Promoter
;
US-09-328-473-1
Query Match          2.3%; Score 175; DB 3; Length 532;
Best Local Similarity 99.3%; Pred. No. 7.9e-57;
Matches 275; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTGGAAGACGTGGTGGAAACGTCCTTTTTCACGGA 6249
Db 289 AGATATCACATCAATCCACTTGTGGAAGACGTGGTGGAAACGTCCTTTTTCACGGA 230
QY 6250 TGTTCCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTGAACGA 6309
Db 229 TGTTCCTCGTGGGTGGGTCCATCTTTGGGACCACTGTGCGTAGAGGCATCTTGAACGA 170
QY 6310 TAGCCTTTCCCTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTCTACTGTCC 6369
Db 169 TAGCCTTTCCCTTATCGCAATGATGGCATTTGTAGAAGCCATCTTCTTCTACTGTCC 110
QY 6370 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429
Db 109 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 50
QY 6430 TTTGTGTAAGTCTCAATAGCCCTCTGCTCTTCTGA 6466
Db 49 TTTGTGTAAGTCTCAATAGCCCTCTGCTCTTCTGA 13
```

```

RESULT 11
US-09-330-737-1/c
; Sequence 1, Application US/09330737
; Patent No. 6232534
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 623253artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/330,737
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,426
; FILING DATE: March 13, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoxie, Thomas
; REGISTRATION NUMBER: 32,993
; REFERENCE/DOCKET NUMBER: 135/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8614
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: 35S Promoter
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-330-737-1

Query Match      2.3%; Score 175; DB 3; Length 532;
Best Local Similarity 99.3%; Pred. No. 7.9e-57;
Matches 275; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTCTTTGAAGACGTGGTTGGAACGTCCTTTTTCACGA 6249
DB 289 AGATATCACATCAATCCACTTCTTTGAAGACGTGGTTGGAACGTCCTTTTTCACGA 230

QY 6250 TGTTCCTCGTGGGTGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
DB 229 TGTTCCTCGTGGGTGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 170

QY 6310 TAGCCCTTCTTTATCGCAATGATGCACTTTGTAGAAGCCATCTTCTTTCTACTGTCC 6369
DB 169 TAGCCCTTCTTTATCGCAATGATGCACTTTGTAGAAGCCATCTTCTTTCTACTGTCC 110

QY 6370 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429
DB 109 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 50

QY 6430 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTCTGA 6466
DB 49 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTCTGA 13
```

```

RESULT 12
US-09-329-169-1/c
; Sequence 1, Application US/09329169
; Patent No. 6329575
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: No. 6329575artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/329,169
; FILING DATE: 09-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/042,426
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoxie, Thomas
; REGISTRATION NUMBER: 32,993
; REFERENCE/DOCKET NUMBER: 135/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8614
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
; TYPE: nucleic acid
```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: 35S Promoter
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-329-169-1

Query Match      2.3%; Score 175; DB 3; Length 532;
Best Local Similarity 99.3%; Pred. No. 7.9e-57;
Matches 275; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTCTTTGAAGACGTGGTTGGAACGTCCTTTTTCACGA 6249
DB 289 AGATATCACATCAATCCACTTCTTTGAAGACGTGGTTGGAACGTCCTTTTTCACGA 230

QY 6250 TGTTCCTCGTGGGTGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
DB 229 TGTTCCTCGTGGGTGGGTGCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 170

QY 6310 TAGCCCTTCTTTATCGCAATGATGCACTTTGTAGAAGCCATCTTCTTTCTACTGTCC 6369
DB 169 TAGCCCTTCTTTATCGCAATGATGCACTTTGTAGAAGCCATCTTCTTTCTACTGTCC 110

QY 6370 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429
DB 109 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 50

QY 6430 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTCTGA 6466
DB 49 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTCTGA 13
```

```

RESULT 13
US-09-330-714A-1/c
; Sequence 1, Application US/09330714A
; Patent No. 6342660
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: No. 6342660artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/330,714A
; FILING DATE: 11-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/042,426
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoxie, Thomas
; REGISTRATION NUMBER: 32,993
; REFERENCE/DOCKET NUMBER: 135/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8614
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
```

```
;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: 35S Promoter
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-330-714A-1

Query Match          2.3%; Score 175; DB 3; Length 532;
Best Local Similarity 99.3%; Pred. No. 7.9e-57;
Matches 275; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGAAACGCTCTCTTTTCCACGA 6249
Db 289 AGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGAAACGCTCTCTTTTCCACGA 230
QY 6250 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGCGTAGAGGCATCTTGAACGA 6309
Db 229 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGCGTAGAGGCATCTTGAACGA 170
QY 6310 TAGCCTTTCCCTTTATCGCAATGATGCGCATTTGTAGAAGCCATCTTCTCTACTGTCC 6369
Db 169 TAGCCTTTCCCTTTATCGCAATGATGCGCATTTGTAGAAGCCATCTTCTCTACTGTCC 110
QY 6370 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429
Db 109 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 50
QY 6430 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGA 6466
Db 49 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGA 13

RESULT 14
US-09-328-826-1/c
; Sequence 1, Application US/09328826
; Patent No. 6399860
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6399860artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,826
; FILING DATE: 09-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,426
; FILING DATE: March 13, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoxie, Thomas
; REGISTRATION NUMBER: 32,993
; REFERENCE/DOCKET NUMBER: 135/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8614
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
;

;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: 35S Promoter
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-328-826-1

Query Match          2.3%; Score 175; DB 3; Length 532;
Best Local Similarity 99.3%; Pred. No. 7.9e-57;
Matches 275; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGAAACGCTCTCTTTTCCACGA 6249
Db 289 AGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGAAACGCTCTCTTTTCCACGA 230
QY 6250 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGCGTAGAGGCATCTTGAACGA 6309
Db 229 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGCGTAGAGGCATCTTGAACGA 170
QY 6310 TAGCCTTTCCCTTTATCGCAATGATGCGCATTTGTAGAAGCCATCTTCTCTACTGTCC 6369
Db 169 TAGCCTTTCCCTTTATCGCAATGATGCGCATTTGTAGAAGCCATCTTCTCTACTGTCC 110
QY 6370 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 6429
Db 109 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCGATATTACCC 50
QY 6430 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGA 6466
Db 49 TTTGTTGAAAAGTCTCAATAGCCCTCTGCTCTTCTGA 13

RESULT 15
US-09-289-170-1/c
; Sequence 1, Application US/09289170
; Patent No. 6573438
; GENERAL INFORMATION:
; APPLICANT: Irvin J. Mettler, Paul S. Dietrich, Ralph Sinibaldi
; TITLE OF INVENTION: DNA Construct Containing Bacillus
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6573438artis Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/289,170
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/042,426
; FILING DATE: March 13, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoxie, Thomas
; REGISTRATION NUMBER: 32,993
; REFERENCE/DOCKET NUMBER: 135/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8614
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
;
```

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: 35S Promoter
US-09-289-170-1

Query Match      2.3%; Score 175; DB 3; Length 532;
Best Local Similarity 99.3%; Pred. No. 7.9e-57;
Matches 275; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTTTTCCACGA 6249
Db      |||||||
QY 289 AGATATCACATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTTTTCCACGA 230
Db      |||||||
QY 6250 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 6309
Db      |||||||
QY 229 TGTTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCATCTTGAACGA 170
Db      |||||||
QY 6310 TAGCCTTTCCCTTTATCGCAATGATGSCATTTGTAGAAGCCATCTTCTTCTACTGTCC 6369
Db      |||||||
QY 169 TAGCCTTTCCCTTTATCGCAATGATGSCATTTGTAGAAGCCATCTTCTTCTACTGTCC 110
Db      |||||||
QY 6370 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATTACCC 6429
Db      |||||||
QY 109 TTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATTACCC 50
Db      |||||||
QY 6430 TTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGA 6466
Db      |||||||
QY 49 TTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGA 13
Db      |||||||

Search completed: December 31, 2005, 19:14:53
Job time : 843 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2005, 07:17:49 ; Search time 469 Seconds  
(without alignments)  
8472.606 Million cell updates/sec

Title: US-10-650-249-1  
Perfect score: 7580  
Sequence: 1 agctctataatcaagaga.....gccgcaccgcgggtgagact 7580

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 4172979 seqs, 262114271 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8345958

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications\_NA\_New.\*  
1: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq2.\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3.\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	98	1.3	11667	US-10-508-263-122	Sequence 122, Appl
C 2	62	0.8	332	US-11-038-981A-5	Sequence 5, Appli
C 3	62	0.8	612	US-11-038-981A-1	Sequence 1, Appli
C 4	62	0.8	835	US-10-391-414-10	Sequence 10, Appl
C 5	62	0.8	1273	US-11-038-981A-18	Sequence 18, Appl
C 6	62	0.8	1273	US-11-038-981A-20	Sequence 20, Appl
C 7	62	0.8	1281	US-11-038-981A-19	Sequence 19, Appl
C 8	62	0.8	1935	US-11-038-981A-25	Sequence 25, Appl
C 9	62	0.8	1935	US-11-038-981A-26	Sequence 26, Appl
C 10	62	0.8	1939	US-11-038-981A-24	Sequence 24, Appl
C 11	62	0.8	1963	US-11-038-981A-27	Sequence 27, Appl
C 12	62	0.8	1963	US-11-038-981A-28	Sequence 28, Appl
C 13	62	0.8	1971	US-11-038-981A-29	Sequence 29, Appl
C 14	62	0.8	3469	US-11-192-801-23	Sequence 23, Appl
C 15	62	0.8	3754	US-11-192-801-15	Sequence 15, Appl
C 16	62	0.8	4149	US-11-192-801-13	Sequence 13, Appl
C 17	59	0.8	125	US-11-038-981A-4	Sequence 4, Appli
C 18	59	0.8	1065	US-11-038-981A-15	Sequence 15, Appl
C 19	59	0.8	1065	US-11-038-981A-17	Sequence 17, Appl
C 20	59	0.8	1073	US-11-038-981A-16	Sequence 16, Appl
C 21	59	0.8	1467	US-11-038-981A-21	Sequence 21, Appl
C 22	59	0.8	1467	US-11-038-981A-22	Sequence 22, Appl
C 23	58	0.8	524	US-11-038-981A-6	Sequence 6, Appli

C 24	58	0.8	1459	7	US-11-038-981A-23	Sequence 23, Appl
C 25	43	0.6	524	6	US-10-986-501-81	Sequence 81, Appl
C 26	43	0.6	687	6	US-10-986-501-103	Sequence 103, App
C 27	43	0.6	9808	6	US-10-966-483-43	Sequence 43, Appl
C 28	43	0.6	9808	7	US-11-021-441-33	Sequence 33, Appl
C 29	39	0.5	7142	6	US-10-516-768-22	Sequence 22, Appl
C 30	35	0.5	559	6	US-10-986-501-83	Sequence 83, Appl
C 31	35	0.5	2105	7	US-11-054-385-11	Sequence 11, Appl
C 32	35	0.5	8931	7	US-11-071-651-16	Sequence 16, Appl
C 33	34	0.4	240	7	US-11-038-981A-2	Sequence 2, Appli
C 34	34	0.4	1176	7	US-11-038-981A-11	Sequence 11, Appl
C 35	34	0.4	1184	7	US-11-038-981A-9	Sequence 9, Appli
C 36	34	0.4	1184	7	US-11-038-981A-10	Sequence 10, Appl
C 37	32	0.4	319608	7	US-11-145-703-1	Sequence 1, Appli
C 38	31	0.4	201	6	US-10-995-561-55887	Sequence 55887, A
C 39	31	0.4	201	6	US-10-995-561-56237	Sequence 56237, A
C 40	31	0.4	201	6	US-10-995-561-56446	Sequence 56446, A
C 41	31	0.4	1234	6	US-10-750-185-30828	Sequence 30828, A
C 42	31	0.4	2195	6	US-10-750-185-24652	Sequence 24652, A
C 43	31	0.4	2513	6	US-10-750-185-56860	Sequence 56860, A
C 44	31	0.4	134499	7	US-11-117-187-192	Sequence 192, App
C 45	31	0.4	171936	6	US-10-933-025-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1  
US-10-508-263-122/c  
; Sequence 122, Application US/10508263  
; Publication No. US20050260754A1  
; GENERAL INFORMATION:  
; APPLICANT: BASF Plant Science GmbH  
; TITLE OF INVENTION: Constructs and methods for regulating gene expression  
; FILE REFERENCE: 53262-20085.00  
; CURRENT APPLICATION NUMBER: US/10/508,263  
; NUMBER OF SEQ ID NOS: 126  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 122  
; LENGTH: 11667  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Description of the artificial sequence: suppression  
; OTHER INFORMATION: construct 2 p3300.1-Toc159-GFP-RNAi  
US-10-508-263-122

Query Match	1.3%	Score 98;	DB 6;	Length 11667;
Best Local Similarity	99.3%	Pred. No. 1.1e-27;		
Matches 148;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	6374	GATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATACCTTTG	6433	
Db	8757	GATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATACCTTTG	8698	
QY	6434	TTGAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATCTTGGAGTA	6493	
Db	8697	TTGAAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATCTTGGAGTA	8638	
QY	6494	GACGAGAGTGCTGCTCCACCATTTGG	6522	
Db	8637	GACGAGAGTGCTGCTCCACCATTTGG	8609	

RESULT 2  
US-11-038-981A-5/c  
; Sequence 5, Application US/11038981A  
; Publication No. US20050283856A1  
; GENERAL INFORMATION:  
; APPLICANT: Conner, Timothy W.  
; APPLICANT: Flasiński, Stanisław  
; APPLICANT: Pang, Sheng Z

```
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 5
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Cauliflower mosaic virus
US-11-038-981A-5

Query Match      0.8%; Score 62; DB 7; Length 332;
Best Local Similarity 100.0%; Pred. No. 4.5e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 122 AGATATCATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 63

QY 6250 TG 6251
Db 62 TG 61

RESULT 3
US-11-038-981A-1/c
; Sequence 1, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Plasinski, Stanislaw
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 1
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Cauliflower mosaic virus
US-11-038-981A-1

Query Match      0.8%; Score 62; DB 7; Length 612;
Best Local Similarity 100.0%; Pred. No. 4.2e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 526 AGATATCATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 467

QY 6250 TG 6251
Db 466 TG 465

RESULT 4
US-10-391-414-10/c
; Sequence 10, Application US/10391414
; Publication No. US20050278799A1
; GENERAL INFORMATION:
; APPLICANT: SHINOZAKI, Kazuko
; APPLICANT: KATSURA, Koji
; APPLICANT: ITO, Yusuke
; TITLE OF INVENTION: Stress Induced Promoter Derived From Rice
; FILE REFERENCE: 382.1041
; CURRENT APPLICATION NUMBER: US/10/391,414
```

```
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: JP 2002-377316
; PRIOR FILING DATE: 2002-12-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 835
; TYPE: DNA
; ORGANISM: Cauliflower mosaic virus
US-10-391-414-10

Query Match      0.8%; Score 62; DB 6; Length 835;
Best Local Similarity 100.0%; Pred. No. 4.1e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 744 AGATATCATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 685

QY 6250 TG 6251
Db 684 TG 683

RESULT 5
US-11-038-981A-18/c
; Sequence 18, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Plasinski, Stanislaw
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 18
; LENGTH: 1273
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-18

Query Match      0.8%; Score 62; DB 7; Length 1273;
Best Local Similarity 100.0%; Pred. No. 3.9e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 134 AGATATCATCATCAATCCACTTGCCTTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 75

QY 6250 TG 6251
Db 74 TG 73

RESULT 6
US-11-038-981A-20/c
; Sequence 20, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Plasinski, Stanislaw
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
```

```
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 20
; LENGTH: 1273
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-20

Query Match          0.8%; Score 62; DB 7; Length 1273;
Best Local Similarity 100.0%; Pred. No. 3.9e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGA 6249
Db 898 AGATATCACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGA 839

QY 6250 TG 6251
Db 838 TG 837

RESULT 7
US-11-038-981A-19/c
; Sequence 19, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 19
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-19

Query Match          0.8%; Score 62; DB 7; Length 1281;
Best Local Similarity 100.0%; Pred. No. 3.9e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGA 6249
Db 524 AGATATCACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGA 465

QY 6250 TG 6251
Db 464 TG 463

RESULT 8
US-11-038-981A-25/c
; Sequence 25, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B

; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 25
; LENGTH: 1935
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-25

Query Match          0.8%; Score 62; DB 7; Length 1935;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGA 6249
Db 727 AGATATCACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGA 786

QY 6250 TG 6251
Db 787 TG 788

RESULT 10
US-11-038-981A-24/c
; Sequence 24, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B

; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 26
; LENGTH: 1935
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-26

Query Match          0.8%; Score 62; DB 7; Length 1935;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGA 6249
Db 727 AGATATCACATCAATCCACTTGTCTTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGA 786

QY 6250 TG 6251
Db 787 TG 788
```



```
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 24
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-24
```

```
Query Match          0.8%; Score 62; DB 7; Length 1939;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTGTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 333 AGATATCACATCAATCCACTTGTGTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 274

QY 6250 TG 6251
Db 273 TG 272
```

```
RESULT 11
US-11-038-981A-27/c
; Sequence 27, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 27
; LENGTH: 1963
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-27
```

```
Query Match          0.8%; Score 62; DB 7; Length 1963;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTGTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 822 AGATATCACATCAATCCACTTGTGTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 763

QY 6250 TG 6251
Db 762 TG 761
```

```
RESULT 12
US-11-038-981A-28
; Sequence 28, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
```

```
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 28
; LENGTH: 1963
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-28
```

```
Query Match          0.8%; Score 62; DB 7; Length 1963;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTGTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 905 AGATATCACATCAATCCACTTGTGTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 964

QY 6250 TG 6251
Db 965 TG 966
```

```
RESULT 13
US-11-038-981A-29/c
; Sequence 29, Application US/11038981A
; Publication No. US20050283856A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Flasiński, Stanisław
; APPLICANT: Pang, Sheng Z
; APPLICANT: You, Jinsong
; TITLE OF INVENTION: CHIMERIC PROMOTERS FOR USE IN PLANTS
; FILE REFERENCE: 38-21(51446)B
; CURRENT APPLICATION NUMBER: US/11/038,981A
; CURRENT FILING DATE: 2005-01-20
; PRIOR APPLICATION NUMBER: 60/537,793
; PRIOR FILING DATE: 2004-01-20
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 29
; LENGTH: 1971
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial promoter sequence
US-11-038-981A-29
```

```
Query Match          0.8%; Score 62; DB 7; Length 1971;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTGTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 6249
Db 1135 AGATATCACATCAATCCACTTGTGTTGAAGACGTGGTTGGAACGTCCTCTTTTCCACGA 1076

QY 6250 TG 6251
Db 1075 TG 1074
```

```
RESULT 14
US-11-192-801-23/c
; Sequence 23, Application US/11192801
; Publication No. US20050273882A1
; GENERAL INFORMATION:
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn
```

; CURRENT APPLICATION NUMBER: US/11/192,801  
; PRIOR FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US/10/232,665  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: US/09/377,466  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 3469  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: expression  
; OTHER INFORMATION: cassette  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (25)..(640)  
; OTHER INFORMATION: P-CaMV.35S  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (664)..(734)  
; OTHER INFORMATION: L-Ta.hcb1  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (748)..(1238)  
; OTHER INFORMATION: I-Os.Act1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1241)..(3199)  
; OTHER INFORMATION: Cry3Bb1 variant 11231mv2  
; FEATURE:  
; NAME/KEY: terminator  
; LOCATION: (3217)..(3450)  
; OTHER INFORMATION: T-Ta.hsp17  
; OTHER INFORMATION: T-Ta.hsp17  
US-11-192-801-23

Query Match 0.8%; Score 62; DB 7; Length 3469;  
Best Local Similarity 100.0%; Pred. No. 3.5e-14;  
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTGAGACGTGTTGAGACGTGTTGAGACGTCTCTTTTCCACGA 6249  
|||  
Db 553 AGATATCACATCAATCCACTTGTGAGACGTGTTGAGACGTGTTGAGACGTCTCTTTTCCACGA 494  
|||  
QY 6250 TG 6251  
||  
Db 493 TG 492

## RESULT 15

US-11-192-801-15/c  
; Sequence 15, Application US/11/192801  
; Publication No. US20050273882A1  
; GENERAL INFORMATION:  
; APPLICANT: Romano, Charles P.  
; TITLE OF INVENTION: Improved Expression of Cry3Bb Insecticidal Protein in Plants  
; FILE REFERENCE: 38-21(15304) Cry3Bb Improved Exp. Corn  
; CURRENT APPLICATION NUMBER: US/11/192,801  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: US/10/232,665  
; PRIOR FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: US/09/377,466  
; PRIOR FILING DATE: 1999-08-19  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 3754  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: expression  
; OTHER INFORMATION: cassette

; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (25)..(640)  
; OTHER INFORMATION: P-CaMV.35S  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (669)..(1472)  
; OTHER INFORMATION: I-Zm.Hsp70  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1490)..(3448)  
; OTHER INFORMATION: Cry3Bb1 variant v11231  
; FEATURE:  
; NAME/KEY: terminator  
; LOCATION: (3475)..(3730)  
; OTHER INFORMATION: Agrobacterium tumefaciens nos 3' transcription  
; OTHER INFORMATION: termination and polyadenylation sequence  
US-11-192-801-15

Query Match 0.8%; Score 62; DB 7; Length 3754;  
Best Local Similarity 100.0%; Pred. No. 3.4e-14;  
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6190 AGATATCACATCAATCCACTTGTGAGACGTGTTGAGACGTGTTGAGACGTCTCTTTTCCACGA 6249  
|||  
Db 553 AGATATCACATCAATCCACTTGTGAGACGTGTTGAGACGTGTTGAGACGTCTCTTTTCCACGA 494  
|||  
QY 6250 TG 6251  
||  
Db 493 TG 492

Search completed: December 31, 2005, 20:22:35  
Job time : 470 secs

**BEST AVAILABLE COPY**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 31, 2005, 05:24:09 ; Search time 17827 Seconds  
(without alignments)  
19893.761 Million cell updates/sec

Title: US-10-650-249-1  
Perfect score: 7580  
Sequence: 1 agctctattaattcaagaga.....gcccaccgcggtgagact 7580

Scoring table: OLIGO\_NUC

Word size: 60.0 ; Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size: 60.0

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hc:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gsa1:\*  
10: gb\_gsa2:\*  
11: gb\_gsa3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	726	9.6	1466	4	CNS0A60U
2	717	9.5	736	10	CW802337
3	711	9.4	720	10	CW797151
4	708	9.3	1371	4	CNS0A5T8
5	580	7.7	868	8	DR749973
6	580	7.7	972	8	DR750042
7	528	7.0	528	10	CNS00PNT
8	411	5.4	866	8	DR750041
9	403	5.3	486	9	B26529
10	382	5.0	433	9	B97814
11	369	4.9	452	9	AQ011657
12	323	4.3	441	7	CK437874
13	273	3.6	446	1	AA585745
14	98	1.3	199	9	AQ361597
15	98	1.3	492	9	AQ361936
16	96	1.3	468	10	CU517384
17	96	1.3	690	7	CF854584
18	96	1.3	839	9	AQ362115
19	89	1.2	464	9	BH173054
20	82	1.1	250	8	T22381
21	81	1.1	378	3	BP861585
22	81	1.1	572	9	CC966736

C 23	81	1.1	761	10	CW800801
C 24	81	1.1	869	10	CW799882
C 25	81	1.1	1320	4	CNS0A504
26	70	0.9	654	10	CG206846
27	63	0.8	735	10	CZ910009
28	62	0.8	115	9	BH751611
29	62	0.8	158	9	BH748291
C 30	62	0.8	175	9	BH746616
31	62	0.8	198	10	CG709685
32	62	0.8	207	9	BZ587759
33	62	0.8	210	9	BH638795
C 34	62	0.8	215	9	BH753813
C 35	62	0.8	219	9	BH747744
C 36	62	0.8	220	9	BH747438
C 37	62	0.8	230	9	BH746474
C 38	62	0.8	242	9	BH638950
C 39	62	0.8	244	9	BH746375
C 40	62	0.8	248	9	BH254798
C 41	62	0.8	251	9	BH750171
C 42	62	0.8	254	9	BH748500
C 43	62	0.8	256	9	BH748499
C 44	62	0.8	268	9	BH611919
45	62	0.8	269	9	BH211646

#### ALIGNMENTS

RESULT 1  
CNS0A60U  
LOCUS  
DEFINITION  
CNS0A60U 1466 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTFB52B04 of Flowers and buds of strain col-0 of Arabidopsis  
thaliana (thale cress).  
ACCESSION  
BX822604.1 GI:42464466  
VERSION  
HTC; GSLT cDNA  
KEYWORDS  
Arabidopsis thaliana (thale cress)  
SOURCE  
Arabidopsis thaliana  
ORGANISM  
Arabidopsis thaliana  
REFERENCE  
1 (bases 1 to 1466)  
Castelli V., Aury J.M., Jaillon O., Wincker P., Clepet C.,  
Menard M., Cruaud C., Quetier P., Scarpelli C., Schachter V.,  
Temple G., Caboche M., Weissenbach J. and Salanoubat M.  
TITLE  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
Unpublished  
2 (bases 1 to 1466)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr]  
COMMENT  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_FF/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
Location/Qualifiers  
1..1466  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"

FEATURES  
Source

gene		/db_xref="taxon:3702" /clone="GSLTFFB52ZB04" /tissue_type="Flowers and buds" /ecotype="Col-0" /plasmid="pCMVSPORT_6" 1..1466 /gene="At3g55370"	
ORIGIN			
Query Match		9.6%; Score 726; DB 4; Length 1466;	
Best Local Similarity		99.5%; Pred. No. 0;	
Matches 1046; Conservative		0; Mismatches 4; Indels 1; Gaps 1;	
QY	4340	CTTACTCAACCTCGCCATTTTCGCAAAACATCTCGTCGTTATTTGACACGTCGGCGGTCC	4399
Db	408	CTTACTCAACCTCGCCATTTTCGCAAAACATCTCGTCGTTATTTGACACGTCGGCGGTCC	467
QY	4400	TTGAGGAATGTTCCCTGTTGGAGGAGCTTTAGGAGGAACAAGAGCAAAATCCAGATCG	4459
Db	468	TTGAGGAATGTTCCCTGTTGGAGGAGCTTTAGGAGGAACAAGATGAAGCAATCCAGATCG	527
QY	4460	AAATCTACGGTCTGCTGCTGATATCTACTAGTACTTCACTTCACTTCTCGC	4519
Db	528	AAATCTACGGTCTGCTGCTGATATCTACTAGTACTTCACTTCACTTCTCGC	587
QY	4520	CCAAGTTACTCAAAACCTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAATCC	4579
Db	588	CCAAGTTACTCAAAACCTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAATCC	647
QY	4580	AACTTCCCATCTTGCTCTCTCCAAAGCCTTGGAGATTACAAATTCAGACCAACTGGA	4639
Db	648	AACTTCCCATCTTGCTCTCTCCAAAGCCTTGGAGATTACAAATTCAGACCAACTGGA	707
QY	4640	TTAGATTTTGGTGGAACTCAAAATAGCAACATGATAGTGGTATGAGTTCTAGTGGTGG	4699
Db	708	TTAGATTTTGGTGGAACTCAAAATAGCAACATGATAGTGGTATGAGTTCTAGTGGTGG	767
QY	4700	ATCTGGATGCATGGAGATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTGATC	4759
Db	768	ATCTGGATGCATGGAGATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTGATC	827
QY	4760	AACTACCGGATTTGGTGGAACTCTCAAAAGCCTTATATCCATTTACTAGAAGTAAGGA	4819
Db	828	AAACTACCGGATTTGGTGGAACTCTCAAAAGCCTTATATCCATTTACTAGAAGTAAGGA	887
QY	4820	GGTGTTAATCAAGGTGATTTCTCAACAAGAGTAGTATTAATTCCAATCAGCTAATGTTT	4879
Db	888	GGTGTTAATCAAGGTGATTTCTCAACAAGAGTAGTATTAATTCCAATCAGCTAATGTTT	947
QY	4880	AGCCCTTGATGATTTTCTTCAGCGGGGTTAGCCCAACCAAGAAATGTGAAG	4939
Db	948	AGCCCTTGATGATTTTCTTCAGCGGGGTTAGCCCAACCAAGAAATGTGAAG	1007
QY	4940	CGGGAAGAAATGATCAGATCGGGTGGAGTGGGATGAGTGAATAACTTATCAAGA	4999
Db	1008	CGGGAAGAAATGATCAGATCGGGTGGAGTGGGATGAGTGAATAACTTATCAAGA	1067
QY	5000	AACTTTTGGGTAATATCAACATAAACTCAGGCAGGAACGAGGAATACATCATCGGGA	5059
Db	1068	AACTTTTGGGTAATATCAACATAAACTCAGGCAGGAACGAGGAATACATCATCGGGA	1127
QY	5060	GGTAACAGTTCTTGGACCGGTTTCACTCCAAACACTCAACAGGCCATCTCTCATTTAA	5119
Db	1128	GGTAACAGTTCTTGGACCGGTTTCACTCCAAACACTCAACAGGCCATCTCTCATTTAA	1186
QY	5120	GTACTCAGCACTAGCTATTTCTTCATGATTTCTTTTGGTGGGTTGATCATTTGGTCTT	5179
Db	1187	GTACTCAGCACTAGCTATTTCTTCATGATTTCTTTTGGTGGGTTGATCATTTGGTCTT	1246
QY	5180	GTCAATCGAGTTATTTGCTGAGGAAGATCAAAACCATGAGCTATATCAAAAGGCTAATTT	5239
Db	1247	GTCAATCGAGTTATTTGCTGAGGAAGATCAAAACCATGAGCTATATCAAAAGGCTAATTT	1306

QY	5240	GAGGCTCAAGGAAGGTATCGTTATATAAACTATCTTTTGGATCTTTTAAAGATCTTCA	5299
Db	1307	GAGGCTCAAGGAAGGTATCGTTATATAAACTATCTTTTGGATCTTTTAAAGATCTTCA	1366
QY	5300	AAGTGTGAGTATGTTTATTTGGTTCCTCGTGTATATTTATGTTTATTAAGAAATTCGT	5359
Db	1367	AAGTGTGAGTATGTTTATTTGGTTCCTCGTGTATATTTATGTTTATTAAGAAATTCGT	1426
QY	5360	CTTATATATTTGGCTATATATAGAGGTGGG	5390
Db	1427	CTTATATATTTGGCTATATATAGAGGTGGG	1457
RESULT 2			
LOCUS	CW802337	736 bp	DNA linear GSS 23-NOV-2004
DEFINITION	Wisconsin477-480J24 Arabidopsis thaliana T-DNA insertion flanking sequences Arabidopsis thaliana genomic, genomic survey sequence.		
ACCESSION	CW802337		
VERSION	CW802337.1		
KEYWORDS	GSS.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
	1 (bases 1 to 736)		
	Woody, S.T., Monson, S.S., Jester, P.J., Austin-Phillips, S., Amasino, R.M., Subesman, M.R. and Krysan, P.J. Small Genes and A New Community Resource for Knock-Out Genes and Translational-Duplicated Gene Families and for Mosaic Analysis in Arabidopsis		
JOURNAL	Unpublished (2004)		
	Contact: Woody ST		
	Biotechnology Center		
COMMENT	University of Wisconsin-Madison		
	425 Henry Mall, Madison, WI 53706, USA		
	Tel: (608) 262-4640		
FEATURES	Email: swoody@facstaff.wisc.edu		
	Class: TAIL-PCR.		
	Location/Qualifiers		
source	1..736		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/cultivar="Col-O ecotype"		
	/db_xref="taxon:3702"		
	/tissue_type="seeds produced by primary (Basta-resistant) transformants"		
	/clone_lib="Arabidopsis thaliana T-DNA insertion flanking sequences"		
	/note="Vector: pDS-Lox; Sequence generated in the course of an Arabidopsis T-DNA tagging program. TAIL-PCR was used to generate sequencing templates that represent A.T. genomic DNA flanking the left border of the pDS-Lox T-DNA insert. PCR products were sequenced directly by using the p745 primer 5' AACGTCGCAATGTTATTAAAGTTGTC 3'"		
	ORIGIN		
Query Match		9.5%; Score 717; DB 10; Length 736;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 717; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	537	GCCTTTCTGGAATAATGCGCATGTTCTTTGGTAATCTAGGAATCTCTCTCTACTAAAG	596
Db	20	GCCTTTCTGGAATAATGCGCATGTTCTTTGGTAATCTAGGAATCTCTCTCTACTAAAG	79
QY	597	ATTGGCATCAGCTGTGAACGATTTCCATTTGTATATCTATCTCGATATTTCCACCTAA	656
Db	80	ATTGGCATCAGCTGTGAACGATTTCCATTTGTATATCTATCTCGATATTTCCACCTAA	139
QY	657	ATATCTTGAATAATTAAGATCAAAATATATATGAGAACATATTTGATCATTTGTTGAATA	716
Db	140	ATATCTTGAATAATTAAGATCAAAATATATATGAGAACATATTTGATCATTTGTTGAATA	199



ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 1371)  
 AUTHORS Castellani, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.  
 TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1371)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castellani V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full\\_length/](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length/)  
<http://www.genoscope.cns.fr/cgi-bin/ggb?ggb?source=Arabidopsis>.

## FEATURES

source  
 1..1371  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="GSLTFB32803"  
 /tissue\_type="Flowers and buds"  
 /ecotype="Col-0"  
 /plasmid="pCWSPORT\_6"  
 1..1371  
 /gene="At3g55370"

ORIGIN  
 Query Match 9.3%; Score 708; DB 4; Length 1371;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4105 GCAGCAGGGAACCAACATCAGTAGAATGTGTCAACACTGACCAACCCCTAATAATTA 4164  
 DB 151 GCAGCAGGGAACCAACATCAGTAGAATGTGTCAACACTGACCAACCCCTAATAATTA 210

QY 4165 CTTACGGCAGCTCTCATCACCCGACTTCTCAGGTTGCAGGTTTCAGGTCAGGTCAGT 4224  
 DB 211 CTTACGGCAGCTCTCATCACCCGACTTCTCAGGTTGCAGGTTTCAGGTCAGT 270

QY 4225 GAATTCATATGGTGGAAAGTGCTCGGATCGCAAAAGTCCCATTCGCTGAAGCAGCTTAAA 4284  
 DB 271 GAATTCATATGGTGGAAAGTGCTCGGATCGCAAAAGTCCCATTCGCTGAAGCAGCTTAAA 330

QY 4285 TTGCCCTAGATGTGACTCAACCAATCAAGTTCGTCTTCTTACTTCAATAATATAGCCTTAC 4344  
 DB 331 TTGCCCTAGATGTGACTCAACCAATCAAGTTCGTCTTCTTACTTCAATAATATAGCCTTAC 390

QY 4345 TCAACCTCGCATTTCTGCAGAAACATGTCTGCTCTATTGGACACGCGGGTTCCTTTGAG 4404  
 DB 391 TCAACCTCGCATTTCTGCAGAAACATGTCTGCTCTATTGGACACGCGGGTTCCTTTGAG 450

QY 4405 GAATGTTCTGTTGGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATC 4464  
 DB 451 GAATGTTCTGTTGGAGGAGGCTTTAGGAGGAAACAAGAGCAAAATCCAGATCGAAATC 510

QY 4465 TAGCGTGTGGTCTCGACTGATATATCTACTAGTACTTCTATCATCTTCTCTCGCCCAAG 4524

DB 511 TACGGTGTGTCTCGACTGATATATCTACTAGTACTTCTATCCTTCTCTCGCCCAAG 570  
 QY 4525 TTACTCAAAACCTTAGCAAGTTTCTAGCTACGGTCAAAATCCGGAGTTTAATTCCAACTT 4584  
 DB 571 TTACTCAAAACCTTAGCAAGTTTCTAGCTACGGTCAAAATCCGGAGTTTAATTCCAACTT 630  
 QY 4585 GCCCATCTTGCTCTCTCTCAAAAGCCTTGGAGATTACAAATTCAGAACACACTGGATTAGA 4644  
 DB 631 GCCCATCTTGCTCTCTCTCAAAAGCCTTGGAGATTACAAATTCAGAACACACTGGATTAGA 690  
 QY 4645 TTTTGGTGGAACTCAAAATPAAGCAACATGATGAAGTGTATGATGTTCTAGTGTGGATCTT 4704  
 DB 691 TTTTGGTGGAACTCAAAATPAAGCAACATGATGAAGTGTATGATGTTCTAGTGTGGATCTT 750  
 QY 4705 GGATGATGGAGATACCTCCATCACACACAGCTCAGCAATTCCTTCTTGTATCAACAC 4764  
 DB 751 GGATGATGGAGATACCTCCATCACACACAGCTCAGCAATTCCTTCTTGTATCAACAC 810  
 QY 4765 TACCGGATTGTGCAATCTTCAACGCGTTATATATCCATTACTTAGAAGG 4812  
 DB 811 TACCGGATTGTGCAATCTTCAACGCGTTATATATCCATTACTTAGAAGG 858

RESULT 5  
 DR749973 868 bp mRNA linear EST 19-JUL-2005  
 LOCUS 79-L025823-045-006-G10-SeqA MF1Z-ADIS-065d Arabidopsis thaliana  
 DEFINITION cDNA clone 006-G10, mRNA sequence.  
 ACCESSION DR749973  
 VERSION DR749973.1 GI:71035313  
 KEYWORDS EST  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 868)  
 AUTHORS Paz-Ares, J., Valencia, A., Costantino, P., Vittorioso, P., Davies, B., Gil-Martin, P., Giraudat, J., Parcy, F., Reindl, A., Sablowski, R., Coupland, G., Martin, C., Angenot, G.C., Baumelein, H., Mock, H.P., Carbonero, P., Colombo, L., Tonelli, C., Engstrom, P., Droege-Laser, W., Gatz, C., Kavanagh, T., Kushnir, S., Zabeau, M., Laux, T., Hordsworth, M., Ruberti, I., Ratcliff, F., Smeekens, S., Somsich, I., Weisshaar, B. and Traas, J.  
 TITLE REGIA, an EU project on functional genomics of transcription factors from Arabidopsis thaliana  
 JOURNAL Comp. Funct. Genomics 3 (2), 102-108 (2002)  
 COMMENT Contact: Paz-Ares, Costantino, Vittorioso, Davies, Gil-Martin, Giraudat, Parcy, Sablowski, Coupland, Martin, Angenot, Baumelein, Carbonero, Colombo, Tonelli, Engstrom, Droege-Laser, Gatz, Kavanagh, Kushnir, Zabeau, Laux, Hordsworth, Ruberti, Smeekens, Somsich, Weisshaar, Traas  
 Bielefeld University, Institute for Genome Research  
 Universitaetsstrasse 25, D-33594 Bielefeld, Germany  
 Email: bernd.weisshaar@uni-bielefeld.de  
 AGI: AT3G55370; Segnalys: truncated in 5' at pos 179;  
 Translation: no full cds detected  
 Data analysis performed in the frame of REGULATORS (Exploiting inter-species conservation in promoter sequences to identify regulators of reproductive development and physiological performance), a Trilateral Co-Operation in Plant Genomics between Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent Thareau (ThP-Orsay UMR818 CNRS-UPS, thareau-ad-ibp.u-psud.fr) and Alain Lecharny (URGV-Evry UMR INRA-CNRS-UEVE, lecharny-ad-ibp.u-psud.fr).  
 Definition of the terms used to describe the quality of the clone: The about 2250 sequences from the clone collection were sorted according to clones and clustered. If more than one contig was formed, the clone was designated 'contamination'. The contigs and singletons were blasted against CDS plus pseudogenes from the TIGR5 annotation, and the resulting AGI code is presented if more



than 90 percent identity was found. The sequences were also blasted against all TIGR5 introns, and matches longer than 50 bp with 95 percent identity are reported as 'intron found'. The remaining terms for SeqAnalysis describe the outcome of the evaluation of the CDS detected after pairwise alignment with CDS plus pseudogenes from the TIGR5 annotation file. The sequences or contigs for which a full CDS with or without STOP codon was detected, a BLASTp against all TIGR5 protein sequences was performed. Full perfect: 100 percent identity; full good: better than 95 percent identity over more than 95 percent of the sequence; partial good: better than 95 percent identity over less than 95 percent of the sequence; weak similarity: less than 95 percent identity over less than 95 percent of the sequence; no similarity: no hit from BLASTp. Note that the collection contains a few clones for which sequencing was not successful, which could have simple technical reasons. Obviously, information about these clones is missing in the submitted data.

Insert Length: 868 Std Error: 0.00  
Seq primer: Seta TCGGTTAACGCTAGCATGGATCTC.

#### FEATURES

Location/Qualifiers  
1. .868  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecolyde="Columbia"  
/db\_xref="GABI:1452143"  
/db\_xref="taxon:3702"  
/clone="006-G10"  
/lab\_host="E. coli DH5alpha"  
/clone\_lib="MP1Z-ADIS-065d"  
/note="Vector: pDONR201; In the context of the EU-funded project REGIA (QLG-CT11999-00876, coordinator Javier Paz-Ares), a set of transcription factor ORFs was generated. The ORFs were produced in a decentralized way in the labs of the participants. Most of the ORFs were generated by RT-PCR using cDNA from various A. thaliana tissues as a template. Initially, it was planned to use yeast recombination to move the ORFs from the cloning vectors into target constructs. For this reason, a number of the clones contain 'REGIA tags' (RG\_tag1:  
5p-RATTCCAGTCACACC-3p; RG\_tag2:  
5p-RATCGCAATCCGGGATC-3p). During the lifetime of the project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequenceing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weishaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by GabiPD (<http://gabi.rzpd.de>)."

#### ORIGIN

Query Match 7.7%; Score 580; DB 8; Length 868;  
Best Local Similarity 100.0%; Pred. No. 1.2e-258;  
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
4233 TGGTGAAGCTGTCGGATCGCAAAAGTCGATCCATTCGTCGAAGCAGCTCTAAATTGGCCCTA 4292  
73 TGGTGAAGCTGTCGGATCGCAAAAGTCGATCCATTCGTCGAAGCAGCTCTAAATTGGCCCTA 132  
4293 GATGTGACTCAACCAATAGTCTGTGTACTTCAATTAACATAGCCCTTACTCAACCTC 4352  
133 GATGTGACTCAACCAATAGTCTGTGTACTTCAATTAACATAGCCCTTACTCAACCTC 192  
4353 GCATTTCTGC AAAACATGTCGTCGCTATTGGACACGTCGGCTTCCTTTCGAGGAATGTTTC 4412  
193 GCATTTCTGC AAAACATGTCGTCGCTATTGGACACGTCGGCTTCCTTTCGAGGAATGTTTC 252  
4413 CTGTTGGAGGAGCTTTAGAGGAAACAAGAGAACAAATCCAGATCGAAATCTACGGTGC 4472  
253 CTGTTGGAGGAGCTTTAGAGGAAACAAGAGAACAAATCCAGATCGAAATCTACGGTGC 312

4473 TGGTCTCGACTGATATATCTACTACTAGTACTTCACTTCACTTCTTCCGCCAAGTTACTCAA 4532  
313 TGGTCTCGACTGATATATCTACTACTAGTACTTCACTTCTTCCGCCAAGTTACTCAA 372  
4533 ACCCTAGCAAGTTTTCATAGCTACGGTCAAAATCCGGAGTTTAAATTCCAACTTCCCATCT 4592  
373 ACCCTAGCAAGTTTTCATAGCTACGGTCAAAATCCGGAGTTTAAATTCCAACTTCCCATCT 432  
4593 TGCCTCTCTCCAAAGCCCTTGAGATTACAATTCAAAGCAACTGGATTAGATTTGGTG 4652  
433 TGCCTCTCTCCAAAGCCCTTGAGATTACAATTCAAAGCAACTGGATTAGATTTGGTG 492  
4653 GAACTCAATATAGCAACATGATAGTGTATAGTCTTAGTGGTGGATCTTGGATGTCAT 4712  
493 GAACTCAATATAGCAACATGATAGTGTATAGTCTTAGTGGTGGATCTTGGATGTCAT 552  
4713 GGAGAAATACCTCCATCACAAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGAT 4772  
553 GGAGAAATACCTCCATCACAAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGAT 612  
4773 TGGTGAATCTTCCAAAGCCCTTATATCCATTACTAGAAG 4812  
613 TGGTGAATCTTCCAAAGCCCTTATATCCATTACTAGAAG 652  
RESULT 6  
DR750042/c  
LOCUS  
DEFINITION  
79-1021445-065-006-G10-SeqLB MP1Z-ADIS-065d Arabidopsis thaliana  
CDNA clone 006-G10, mRNA sequence.  
ACCESSION  
DR750042  
VERSION  
DR750042.1 GI:71035382  
KEYWORDS  
EST.  
SOURCE  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 972)  
Paz-Ares, J., Valencia, A., Costantino, P., Vittorioso, P., Davies, B.,  
Gilmarin, P., Giraudat, J., Parcy, E., Reindl, A., Sablowski, R.,  
Coulpland, G., Martin, C., Angenent, G. C., Baumelein, H., Mock, H. P.,  
Carbonero, P., Colombo, L., Tonelli, C., Engstroem, P.,  
Droge-Laser, W., Gatz, C., Kavanagh, T., Kushnir, S., Zabeau, M.,  
Lauz, T., Hordsworth, M., Ruberti, I., Ratcliff, F., Smeekens, S.,  
Somsich, I., Weishaar, B. and Traas, J.  
REGIA, an EU project on functional genomics of transcription  
factors from Arabidopsis thaliana  
Comp. Funct. Genomics 3 (2), 102-108 (2002)  
Contact:  
Paz-Ares, Costantino, Vittorioso, Davies, Gilmarin, Giraudat, Parcy, Sabl  
owski, Coulpland, Martin, Angenent, Baumelein, Carbonero, Colombo, Tonelli,  
Engstroem, Droge-Laser, Gatz, Kavanagh, Kushnir, Zabeau, Lauz, Hordsworth  
Ruberti, Smeekens, Somsich, Weishaar, Traas  
Bielefeld University, Institute for Genome Research  
Universitaetsstrasse 25, D-33594 Bielefeld, Germany  
Email: bernd.weishaar@uni-bielefeld.de  
AGI: AT3G53370; SeqAnalysis: truncated in 5' at pos 179;  
Translation: no full cds detected  
Data analysis performed in the frame of REGULATORS (Exploiting  
inter-species conservation in promoter sequences to identify  
regulators of reproductive development and physiological  
performance), a Trilateral Co-Operation in Plant Genomics between  
Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by  
G. Coulpland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent  
Thareau (tbp-orsay UMR818 CNRS-UPS, thareau-ad-ibp.u-psud.fr) and  
Alain Lecharny (URGV-Evry UMR INRA-CNRS-UEVE,  
lecharny-ad-ibp.u-psud.fr).  
Definition of the terms used to describe the quality of the clone:  
The about 2250 sequences from the clone collection were sorted  
according to clones and clustered. If more than one contig was  
formed, the clone was designated 'contamination'. The contigs and



```
Db 61 ACTTCGGGATAAAAATCTCGCATGCAAGGTAATTTATTTTTCATGACAAAAGCCA 120
QY 2127 CAATGGCGATAATTATACATATATACTATGCAAAACGAAACTTTATCTGGGTCAACCG 2186
Db 121 CAATGGCGATAATTATATACTATATACTATGCAAAACGAAACTTTATCTGGGTCAACCG 180
QY 2187 AGGAAACAAAGGTACACTCGATTGAGCACTCCACCAAGACCAACCCCACTACCAAT 2246
Db 181 AGGAAACAAAGGTACACTCGATTGAGCACTCCACCAAGACCAACCCCACTACCAAT 240
QY 2247 CACCTTTATTTGTTCTTTATTCACCTCAAAATCTTTTAAATTTTAAATTTCAAT 2306
Db 241 CACCTTTATTTGTTCTTTATTCACCTCAAAATCTTTTAAATTTTAAATTTCAAT 300
QY 2307 ATTGCTTCTCTCGTTGTTTATTTTAACTTTTAAATTTTAAATTTTAAATTTTAA 2366
Db 301 ATTGCTTCTCTCGTTGTTTATTTTAACTTTTAAATTTTAAATTTTAAATTTTAA 360
QY 2367 TTTCTCTCTGATGCGCACTCTCGCTCTTCACATCTTTTGGGCAACCATTTGTTAC 2426
Db 361 TTTCTCTCTGATGCGCACTCTCGCTCTTCACATCTTTTGGGCAACCATTTGTTAC 420
QY 2427 GTATGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 2486
Db 421 GTATGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 480
QY 2487 TGTGCTTCTCTCATAAAAAGAAATTTATTTAACTTTATTTAGTCCAA 2534
Db 481 TGTGCTTCTCTCATAAAAAGAAATTTATTTAACTTTATTTAGTCCAA 528
```

```
RESULT 8
DR750041
LOCUS
DEFINITION
79-1021444-065-006-G10-SeLA MP1Z-ADIS-065d Arabidopsis thaliana
CDNA clone 006-G10, mRNA sequence.
```

```
ACCESSION
DR750041.1 GI:71035381
VERSION
KEYWORDS
SOURCE
Arabidopsis thaliana (thale cress)
```

```
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
```

```
REFERENCE
1 (bases 1 to 866)
Paz-Ares,J., Valencia,A., Costantino,P., Vittorioso,P., Davies,B.,
Gilmarin,P., Giraudat,J., Parcy,E., Reindl,A., Sablowski,R.,
Coupland,G., Martin,C., Angenent,G.C., Baumelein,H., Mock,H.P.,
Carbonero,P., Colombo,L., Tonelli,C., Engstroem,P.,
Droge-Laser,W., Gatz,C., Kavanagh,T., Kushnir,S., Zabeau,M.,
LauX,T., Hordeworth,M., Ruberti,I., Ratcliff,F., Smeekens,S.,
Somsich,I., Weishaar,B. and Traas,J.
```

```
TITLE
REGIA, an EU project on functional genomics of transcription
factors from Arabidopsis thaliana
Comp. Funct. Genomics 3 (2), 102-108 (2002)
```

```
JOURNAL
COMMENT
Contact:
Paz-Ares, Costantino, Vittorioso, Davies, Gilmarin, Giraudat, Parcy, Sablowski, Coupland, Martin, Angenent, Baumelein, Carbonero, Colombo, Tonelli, Engstroem, Droge-Laser, Gatz, Kavanagh, Kushnir, Zabeau, LauX, Hordeworth, Ruberti, Smeekens, Somsich, Weishaar, Traas,
Bielefeld University, Institute for Genome Research
Universitaetsstrasse 25, D-33594 Bielefeld, Germany
Email: bernd.weishaar@uni-bielefeld.de
```

```
AGI: AT3G55370; SeqAnalysis: truncated in 5' at pos 179;
Translation: no full cds detected
Data analysis performed in the frame of REGULATORS (Exploiting
inter-species conservation in promoter sequences to identify
regulators of reproductive development and physiological
performance), a Trilateral Co-Operation in Plant Genomics between
Spain (MCYT), France (GENOPLANTE) and Germany (GABI) coordinated by
G. Coupland (coupland-ad-mpiz-koeln.mpg.de). Authors: Vincent
Thareau (ThP-Orsay UMR8618 CNRS-UPS, thareau-ad-ibp.u-psud.fr) and
Alain Lechamy (URGV-Evry UMR INRA-CNRS-UEVE,
```

lechamy-ad-ibp.u-psud.fr).  
Definition of the terms used to describe the quality of the clone:  
The about 2250 sequences from the clone collection were sorted according to clones and clustered. If more than one contig was formed, the clone was designated 'Contamination'. The contigs and singletons were blasted against CDS plus pseudogenes from the TIGR5 annotation, and the resulting AGI code is presented if more than 90 percent identity was found. The sequences were also blasted against all TIGR5 introns, and matches longer than 50 bp with 95 percent identity are reported as 'intron found'. The remaining terms for SeqAnalysis describe the outcome of the evaluation of the CDS detected after pairwise alignment with CDS plus pseudogenes from the TIGR5 annotation file. The sequences or contigs for which a full CDS with or without STOP codon was detected, a BLASTp against all TIGR5 protein sequences was performed. Full perfect: 100 percent identity; full good: better than 95 percent identity over more than 95 percent of the sequence; partial good: better than 95 percent identity over less than 95 percent of the sequence; weak similarity: less than 95 percent identity over less than 95 percent of the sequence; no similarity: no hit from BLASTp. Note that the collection contains a few clones for which sequencing was not successful, which could have simple technical reasons. Obviously, information about these clones is missing in the submitted data.

Insert Length: 866 Std Error: 0.00  
Seq primer: SeLA TCGGTTAACGCTAGCATGGATCTC.

FEATURES  
source

Location/Qualifiers

1. .866

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/ecotype="Columbia"

/db\_xref="GABI:1452132"

/db\_xref="taxon:3702"

/clone="006-G10"

/lab\_host="E. coli DH5alpha"

/clone\_lib="MP1Z-ADIS-065d"

/note="Vector: pDONR201; in the context of the EU-funded Project REGIA (QLG-CT11999-00876, coordinator Javier Paz-Ares), a set of transcription factor ORFs was generated. The ORFs were produced in a decentralized way in the labs of the participants. Most of the ORFs were generated by RT-PCR using cDNA from various A. thaliana tissues as a template. Initially, it was planned to use yeast recombination to move the ORFs from the cloning vectors into target constructs. For this reason, a number of the clones contain 'REGIA tags' (RG\_tag1: 5pr-AATTCCAGCTCACCC-3pr; RG\_tag2: 5pr-CATGCGAATTCCTCCGGGATC-3pr). During the lifetime of the project, the GATEWAY system became available and finally all ORFs were transferred into GATEWAY vectors. At the end of the REGIA project, the clones were collected as plasmid DNA. For end-sequencing at the DNA core facility of the MPI of Plant Breeding Research (ADIS; head: Bernd Weishaar), the plasmids were re-transformed into DH5alpha, arrayed into 96-well plates and new plasmid DNA was prepared. Re-transformation proved to be difficult because a number of clones displayed poor growth. Data submission has been handled by Gabipd (<http://gabi.rzpd.de>)."

ORIGIN

Query Match 5.4%; Score 411; DB 8; Length 866;  
Best Local Similarity 99.5%; Pred. No. 8.6e-180;  
Matches 561; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4249 GATCCAAAAGTCCCATTCCTCGAAGCAGCTCTAAATTCCTAGATGACTCAACCA 4308

Db 34 GATCCAAAAGTCCCATTCCTCGAAGCAGCTCTAAATTCCTAGATGACTCAACCA 93

QY 4309 TACTAAGTTCTGTACTTCAATACTATAGCCTTACTCAACCTCGCCATTTCTGCAAAAC 4368

Db 94 TACTAAGTTCTGTACTTCAATACTATAGCCTTACTCAACCTCGCCATTTCTGCAAAAC 153

```
QY 4369 ATGTCGTGCTATTGACACAGCGGGGCTTCTCTGAGGAATGTTCTGTTGGAGGAGGCTT 4428
|||||
Db 154 ATGTCGTGCTATTGACACAGCGGGGCTTCTCTGAGGAATGTTCTGTTGGAGGAGGCTT 213
|||||
QY 4429 TAGGAGGAACAAGAGAGCAAAATCCAGATCGAAATCTACGGTCGTGGTCTCGACTGATAA 4488
|||||
Db 214 TAGGAGGAACAAGAGAGCAAAATCCAGATCGAAATCTACGGTCGTGGTCTCGACTGATAA 273
|||||
QY 4489 TACTACTAGTACTTCACTCACTTACTCTCGCCCAAGTTACTCAAAACCTCAGCAAGTTTCA 4548
|||||
Db 274 TACTACTAGTACTTCACTCACTTACTCTCGCCCAAGTTACTCAAAATTTAGCAAGTTTCA 333
|||||
QY 4549 TAGCTACGCTCAAAATCCCGGAGTTTAAATTCCAACTTGGCCATCTTGCCTCTCCAAAG 4608
|||||
Db 334 TAGCTACGCTCAAAATCCCGGAGTTTAAATTCCAACTTGGCCATCTTGCCTCTCCAAAG 393
|||||
QY 4609 CTTTGGAGATTACAATTCGAACAACACATGGATTAGATTGTTGGTGAACCTCAAAATAGCAA 4668
|||||
Db 394 CTTTGGAGATTACAATTCGAACAACACATGGATTAGATTGTTGGTGAACCTCAAAATAGCAA 453
|||||
QY 4669 CATGATAAGTGTATGAGTTCTAGTGGTGGGATCTTGGATGATGGAGATACCTCCATC 4728
|||||
Db 454 CATGATAAGTGTATGAGTTCTAGTGGTGGGATCTTGGATGATGGAGATACCTCCATC 513
|||||
QY 4729 ACAACAAGCTCAGCAATCCCTTTCTTGATCAACACTACCGGATTTGGTGCATCTTCAA 4788
|||||
Db 514 ACAACAAGCTCAGCAATCCCTTTCTTGATCAACACTACCGGATTTGGTGCATCTTCAA 573
|||||
QY 4789 CGCGTTATATCCATTACTAGAGG 4812
|||||
Db 574 CGCGTTATATCCATTACTAGAGG 597
|||||
```

```
RESULT 9
B26529
LOCUS F9G12TF IGF Arabidopsis thaliana genomic clone F9G12, genomic
DEFINITION survey sequence.
ACCESSION B26529
VERSION B26529.1 GI:2512495
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 486)
AUTHORS Rounsley,S.D., Kelley,J.M., Field,C.E., Craven,M.B., Adams,M.D. and
Venter,J.C.
TITLE Use of a BAC End Sequence Database To Identify Minimal Overlaps for
Arabidopsis Genomic Sequencing
JOURNAL Unpublished (1997)
COMMENT Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends
High quality sequence stop: 486.
```

```
Location/Qualifiers
1. .486
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="F9G12"
/sex="hermaphrodite"
/clone_lib="IGF"
/note="Vector: BelOBA2II; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"
```

```
ORIGIN
Query Match 5.3%; Score 403; DB 9; Length 486;
Best Local Similarity 100.0%; Pred. No. 4.9e-176;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2031 TCTTAAACAGAGGAATTTTACATCATCTTCTAGACTGAACTTTTCGGGATAAAAAATCTCGCC 2090
|||||
Db 8 TCTTAAACAGAGGAATTTTACATCATCTTCTAGACTGAACTTTTCGGGATAAAAAATCTCGCC 67
|||||
QY 2091 ATGCAAGAGTAAATTTATTTTTCATGACAAAGCCACAATGCGCGATATTAATACATATAA 2150
|||||
Db 68 ATGCAAGAGTAAATTTATTTTTCATGACAAAGCCACAATGCGCGATATTAATACATATAA 127
|||||
QY 2151 TACTATGCAAAACGAACTTTTACTTGGGTCTATACCGAGGAAAAACAAGGTACACTCGATT 2210
|||||
Db 128 TACTATGCAAAACGAACTTTTACTTGGGTCTATACCGAGGAAAAACAAGGTACACTCGATT 187
|||||
QY 2211 GTGCAAACTCCACCAAGACCAACCCACTACCAATTCACCTTATTTGTTCTTTATTCA 2270
|||||
Db 188 GTGCAAACTCCACCAAGACCAACCCACTACCAATTCACCTTATTTGTTCTTTATTCA 247
|||||
QY 2271 CTCAAAACTCTTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2330
|||||
Db 248 CTCAAAACTCTTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 307
|||||
QY 2331 ACCTTTTAAATTAATAATTTGAAAGGAGGTCCTTAGGTTTCTCTCTCATGCGCCACTCT 2390
|||||
Db 308 ACCTTTTAAATTAATAATTTGAAAGGAGGTCCTTAGGTTTCTCTCTCATGCGCCACTCT 367
|||||
QY 2391 CGCTCTTCACTCTTTTGGGCAACCAATTTGTAACGTATGCA 2433
|||||
Db 368 CGCTCTTCACTCTTTTGGGCAACCAATTTGTAACGTATGCA 410
|||||
```

```
RESULT 10
B97814
LOCUS F18P18TFB IGF Arabidopsis thaliana genomic clone F18P18, genomic
DEFINITION survey sequence.
ACCESSION B97814
VERSION B97814.1 GI:2999893
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 433)
AUTHORS Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and
Venter,J.C.
TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in
Arabidopsis Genomic Sequencing. Update 3
JOURNAL Unpublished (1997)
COMMENT Other GSSs: F18P18TR
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends
High quality sequence stop: 433.
```

```
Location/Qualifiers
1. .433
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="F18P18"
```

```
/sex="hermaphrodite"
/clone_lib="IGF"
/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"

ORIGIN

Query Match      5.0%; Score 382; DB 9; Length 433;
Best Local Similarity 99.8%; Pred. No. 3.1e-166;
Matches 432; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5459 CTCATTTCTTCGAGACCCATTATGACATTCGACATCTATAGAACATATATGTAATG 5518
DB 1 CTCATTTCTTCGAGACCCATTATGACATTCGACATCTATAGAACATATATGTAATG 60

QY 5519 TATATTAACGCTAACTGAAATTTTATGACCAAGTAAATAAATATGCGGAATGTA 5578
DB 61 TATATTAAGGTACTTTAAGTCGAATTTTATGACCAAGTAAATAAATATGCGGAATGTA 120

QY 5579 CATGCTAATATCGAGTTTAACTATTTTCCAATATACAACTATTTCTCTTCGTC 5638
DB 121 CATGCTAATATCGAGTTTAACTATTTTCCAATATACAACTATTTCTCTTCGTC 180

QY 5639 AACTTATATACCTTATTCGATCTTATTTTCTTTTAAATTCCTTTTCTTTCC 5698
DB 181 AACTTATATACCTTATTCGATCTTATTTTCTTTTAAATTCCTTTTCTTTCC 240

QY 5699 AAGACACAAAAAATAACAGAAACGAAAAAAGAGATTTTAAAAAATTCATAACCC 5758
DB 241 AAGACACAAAAAATAACAGAAACGAAAAAAGAGATTTTAAAAAATTCATAACCC 300

QY 5759 ACGAAGATTATGCACTAAATTCAGACTAATCCCAAAATTCAGAAATTTATGATTTT 5818
DB 301 ACGAAGATTATGCACTAAATTCAGACTAATCCCAAAATTCAGAAATTTATGATTTT 360

QY 5819 TCGGATTAATATGTTGTCACAACTAATGCGCAACTAATTCGAAAGACAATGG 5878
DB 361 TCGGATTAATATGAGTTTCAATCAATATGCGCAACTAATTCGAAAGACAATGG 420

QY 5879 AATGACTGAAACC 5891
DB 421 AATGACTGNAACC 433

RESULT 11
AQ011657
LOCUS F24E16T6B IGF Arabidopsis thaliana genomic clone F24E16, genomic
DEFINITION survey sequence.
ACCESSION AQ011657
VERSION AQ011657.1 GI:3166407
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 452)
Rounsley,S.D., Suh,E.J., Wible,C., Golden,K., Shateman,S., Choi,P.,
Yu,K., Akinretoy,B., Shen,K., Goonasekaram,S., Militscher,J.,
Adams,M.D. and Venter,J.C.
A BAC End Sequence Database for Identifying Minimal Overlaps in
Arabidopsis Genomic Sequencing. Update 4
Unpublished (1998)
Other_GSSs: F24E16TR
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends

High quality sequence stop: 452.
Location/Qualifiers
1..452
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="F24E16"
/sex="hermaphrodite"
/clone_lib="IGF"
/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"

ORIGIN

Query Match      4.9%; Score 369; DB 9; Length 452;
Best Local Similarity 99.8%; Pred. No. 3.6e-160;
Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4240 ACGTCTCGGATCGCAAAAGTCCCATTCGCTGAAGCAGCTCTAAATTCGCCCTAGATGTA 4299
DB 12 ACGTCTCGGATCGCAAAAGTCCCATTCGCTGAAGCAGCTCTAAATTCGCCCTAGATGTA 71

QY 4300 CTCACCAATATCTAAGTTCTGTACTTCAATAAATATAGCCCTTACTCAACCTCGCATTT 4359
DB 72 CTCACCAATATCTAAGTTCTGTACTTCAATAAATATAGCCCTTACTCAACCTCGCATTT 131

QY 4360 CTGCAAAACATGTCTGCTATTTGACACGTGGGGTTCCTTTGAGGAATGTTCTGTGG 4419
DB 132 CTGCAAAACATGTCTGCTATTTGACACGTGGGGTTCCTTTGAGGAATGTTCTGTGG 191

QY 4420 AGGAGGCTTTAGGAGGAACAGAGAAATCCAGATCGAATCTACGGTCTGTGCTC 4479
DB 192 AGGAGGCTTTAGGAGGAACAGAGAAATCCAGATCGAATCTACGGTCTGTGCTC 251

QY 4480 GACTCATATATCTACTAGTACTTTCATCTACTTCTCGCCCAAGTACTCAAAACCTAG 4539
DB 252 GACTCATATATCTACTAGTACTTTCATCTACTTCTCGCCCAAGTACTCAAAACCTAG 311

QY 4540 CAAGTTTCATAGTACGTCAAAATCCCGAGTTTAAATTCAACTTGCCTCC 4599
DB 312 CAAGTTTCATAGTACGTCAAAATCCCGAGTTTAAATTCAACTTGCCTCC 371

QY 4600 TCTCCAAAGCCTTGAGATTACAAATTCAGCACTGATAGATTGTTGGTGGACTCA 4659
DB 372 TCTCCAAAGCCTTGAGATTACAAATTCAGCACTGATAGATTGTTGGTGGACTCA 431

RESULT 12
CK437874/c
LOCUS CK437874.1 GI:40768309
DEFINITION Picea glauca (white spruce)
ACCESSION CK437874
VERSION CK437874.1
KEYWORDS EST.
SOURCE Picea glauca (white spruce)
ORGANISM Picea glauca
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
1 (bases 1 to 441)
Morency,M.-J., Cooke,J., Pavy,N., Parsons,L., Paule,C., Seguin,A.,
Retzel,E., Butterfield,Y., Barber,S., Yang,G., Stott,J.,
Siddiqui,A., Holt,R., Marra,M. and Mackay,J.
Arborea EST sequencing in Picea glauca (white spruce)
Unpublished (2004)
Contact: John Mackay
Centre de Recherche en Biologie Forestiere
Universite Laval
Pavillon Charles-Eugene Marchand, Quebec, CANADA G1K 7P4
Fax: 418 656 7493
Email: jmackay@rsvs.ulaval.ca
Center for Computational Genomics and Bioinformatics (CCGB),
```

University of Minnesota, MN id Identifier: MN5176207 Clone ID: GQ0042 P12 Clones available through: John Mackay, Ph. D. Professeur adjoint -Assistant professor EMAIL: jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere (Forest Biology Research Center) Universite Laval Quebec, Quebec CANADA G1K 7P4  
Plate: 2 row: 12 column: p  
Seq primer: M13 Reverse Primer.

## FEATURES

Location/Qualifiers  
1. .441  
/organism="Picea glauca"  
/mol\_type="mRNA"  
/strain="Two trees of provenance 5333 and one from 5206"  
/db\_xref="taxon:3330"  
/clone="GenomeQuebec\_id:GQ0042P12"  
/sex="Hermaphrodite"  
/tissue\_type="Non-lignified differentiating secondary xylem from normal vertical trees"  
/dev\_stage="Non-lignified secondary xylem from trees harvested 2.5 hours, 6 hours and 11 hours after day break"  
/lab\_host="E. coli DH10B cells"  
/clone\_lib="GQ004: Non-lignified secondary xylem from mature trees"  
/note="Organ: Stem from ground to lower part of live crown, on 33 year old tree; Vector: pBluescript II SK (+) XR; Site\_1: Eco-RI; Site\_2: Xho-I; Tissue was harvested in mid-June, during formation of early wood. cDNA was prepared from 5 mg of poly A+ selected RNA and was directionally ligated into the pBluescript II SK (+) XR vector (Stratagene), transformed by electroporation into DH10B cells (In vitro) for propagation"

## ORIGIN

Query Match 4.3%; Score 323; DB 7; Length 441;  
Best Local Similarity 100.0%; Pred. No. 1e-138;  
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6190 AGATATCATCAATCCACTTCTTGAAGAGCTGGTTGGACGCTCTTTTTCACGA 6249  
DB 349 AGATATCATCAATCCACTTCTTGAAGAGCTGGTTGGACGCTCTTTTTCACGA 290  
QY 6250 TGTCTCTGCTGGTGGGCTGCATCTTGGGACCACTGTCGGTAGAGGATCTTGAACGA 6309  
DB 289 TGTCTCTGCTGGTGGGCTGCATCTTGGGACCACTGTCGGTAGAGGATCTTGAACGA 230  
QY 6310 TAGCCCTTCTTTATCGCAATGATGCAATTTGTAGAAGCCATCTTCTCTTCTACTGTC 6369  
DB 229 TAGCCCTTCTTTATCGCAATGATGCAATTTGTAGAAGCCATCTTCTCTTCTACTGTC 170  
QY 6370 TTTGATGAAGTCAGATAGCTGGCAATGGAATCCGAGGAGTTTCCGATATATACC 6429  
DB 169 TTTGATGAAGTCAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATATACC 110  
QY 6430 TTTGTTGAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTGATATCTTGG 6489  
DB 109 TTTGTTGAAAGTCTCAATAGCCCTCTGCTCTCTGAGACTGTATCTTTGATATCTTGG 50  
QY 6490 AGTAGACGAGTGTCTGTCCTCC 6512  
DB 49 AGTAGACGAGTGTCTGTCCTC 27

## RESULT 13

AA585745  
LOCUS 28394 Lambda-PRL2 Arabidopsis thaliana cDNA clone 104F10XP 3', mRNA  
DEFINITION sequence.  
ACCESSION AA585745  
VERSION AA585745.1 GI:2393157  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

## REFERENCE

AUTHORS Newman,T., deBruijn,F.J., Green,P., Keestra,K., Kende,H., McIntosh,L., Ohlrogge,S., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.

## TITLE

Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones

## JOURNAL

Plant Physiol. 106, 1241-1255 (1994)

## PUBMED

7846151

## COMMENT

Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcm@bm.cl.msu.edu

The sequence entry for this EST has been reverse complimented and is being submitted in the sense orientation.

Seq primer: M13 -21 dye primer.

## FEATURES

Location/Qualifiers  
1. .446  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="104F10XP"  
/clone\_lib="Lambda-PRL2"  
/note="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."

## ORIGIN

Query Match 3.6%; Score 273; DB 1; Length 446;  
Best Local Similarity 99.7%; Pred. No. 2.2e-115;  
Matches 323; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 5211 CCATCAGCTATATCCAAAGGCTAATTTGAGGCTCAAGGAAGGTATGTTATAAAC 5270  
DB 123 CCATCAGCTATATCCAAAGGCTAATTTGAGGCTCAAGGAAGGTATGTTATAAAC 182  
QY 5271 TATCTTTTGTATCTTTTAAAGATCTTCAAGTGTGAGTATCTTTATTTGGTTCG 5330  
DB 183 TATCTTTTGTATCTTTTAAAGATCTTCAAGTGTGAGTATCTTTATTTGGTTCG 242  
QY 5331 GTGATATTTATCTTTTATTTAGAAATTTGGTCTTATATATTTGGCTATATATAGAGGTGG 5390  
DB 243 GTGATATTTATCTTTTATTTAGAAATTTGGTCTTATATATTTGGCTATATATAGAGGTGG 302  
QY 5391 TGATATGTATGAATTCAGAGTTGATTTGGAACTTTTTCGTTTCATTTCAATAATC 5450  
DB 303 TGATATGTATGAATTCAGAGTTGATTTGGAACTTTTTCGTTTCATTTCAATAATC 362  
QY 5451 ATCGAATTCCTCAATTTCTTTGGAGACCATTATGAGACATTTAGACATCTATAGAAATAT 5510  
DB 363 AUCGAATTCCTCAATTTCTTTGGAGACCATTATGAGACATTTAGACATCTATAGAAATAT 422  
QY 5511 ATGTAATGTATATTAACGCTACTT 5534  
DB 423 ATGTAATGTATATTAACGCTACTT 446

## RESULT 14

AQ361597/c  
LOCUS AQ361597  
DEFINITION mgx0004F23f CUGI Rice Blast BAC Library Magnaporthe grisea genomic

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.



```

clone mgxb0004F23f, genomic survey sequence.
AQ361597
VERSION AQ361597.1 GI:4211436
KEYWORDS GSS.
SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM Magnaporthe grisea
REFERENCE 1 (bases 1 to 199)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence stop: 96.
FEATURES
source
Location/Qualifiers
1..199
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0004F23f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
ORIGIN
Query Match 1.3%; Score 98; DB 9; Length 199;
Best Local Similarity 99.3%; Pred. No. 1.1e-33;
Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6374 GATGAAGTCACAGATAGCTGGGCAATCGAGAGGTTTCCCGATATTACCCCTTTG 6433
Db 155 GATGAAGTCACAGATAGCTGGGCAATCGAGAGGTTTCCCGATATTACCCCTTTG 96
QY 6434 TTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTTCTGGAGTA 6493
Db 95 TTGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTTCTGGAGTA 36
QY 6494 GACGAGAGTGTGCTGCTCCACCATGTTGG 6522
Db 35 GACGAGAGTGTGCTGCTCCACCATGTTGG 7
Search completed: December 31, 2005, 19:00:45
Job time : 17831 secs

clone mgxb0004F23f, genomic survey sequence.
AQ361597
VERSION AQ361597.1 GI:4211436
KEYWORDS GSS.
SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM Magnaporthe grisea
REFERENCE 1 (bases 1 to 199)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence stop: 96.
FEATURES
source
Location/Qualifiers
1..199
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0004F23f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
ORIGIN
Query Match 1.3%; Score 98; DB 9; Length 199;
Best Local Similarity 99.3%; Pred. No. 1.1e-33;
Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6374 GATGAAGTCACAGATAGCTGGGCAATCGAGAGGTTTCCCGATATTACCCCTTTG 6433
Db 155 GATGAAGTCACAGATAGCTGGGCAATCGAGAGGTTTCCCGATATTACCCCTTTG 96
QY 6434 TTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTTCTGGAGTA 6493
Db 95 TTGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTTCTGGAGTA 36
QY 6494 GACGAGAGTGTGCTGCTCCACCATGTTGG 6522
Db 35 GACGAGAGTGTGCTGCTCCACCATGTTGG 7
Search completed: December 31, 2005, 19:00:45
Job time : 17831 secs

clone mgxb0005M03f, genomic survey sequence.
AQ361936
VERSION AQ361936.1 GI:4211775
KEYWORDS GSS.
SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM Magnaporthe grisea
REFERENCE 1 (bases 1 to 492)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence stop: 75.
FEATURES
source
Location/Qualifiers
1..492
/organism="Magnaporthe grisea"
/mol_type="genomic DNA"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0005M03f"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/clone_lib="CUGI Rice Blast BAC Library"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
ORIGIN
Query Match 1.3%; Score 98; DB 9; Length 492;
Best Local Similarity 99.3%; Pred. No. 9.9e-34;
Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6374 GATGAAGTCACAGATAGCTGGGCAATCGAGAGGTTTCCCGATATTACCCCTTTG 6433
Db 155 GATGAAGTCACAGATAGCTGGGCAATCGAGAGGTTTCCCGATATTACCCCTTTG 96
QY 6434 TTGAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGATATTTCTGGAGTA 6493
Db 95 TTGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTTCTGGAGTA 36
QY 6494 GACGAGAGTGTGCTGCTCCACCATGTTGG 6522
Db 35 GACGAGAGTGTGCTGCTCCACCATGTTGG 7
Search completed: December 31, 2005, 19:00:45
Job time : 17831 secs

```



**BEST AVAILABLE COPY**



[illegible]

[illegible]

Db 4681 TATGATGTTCTAGTGGGATCTTGATGCAATGGAGATACCTCATCAACAAGCTCA 4740  
Qy 4741 GCAATTCCTCTTCTTGATCAACACTACCGGATGGTGAATCTTCAACCGGTTATATCC 4800  
Db 4741 GCAATTCCTCTTCTTGATCAACACTACCGGATGGTGAATCTTCAACCGGTTATATCC 4800  
Qy 4801 ATTACTAGAGGTAGAGGAGGTGTTAATCAAGGTGATTTCTCAACAGAGAGTAGTGATTA 4860  
Db 4801 ATTACTAGAGGTAGAGGAGGTGTTAATCAAGGTGATTTCTCAACAGAGAGTAGTGATTA 4860  
Qy 4861 TTCCCAATCAGCTAATGTTTAAAGCCCTTGATGATTTTCTCAGCGGGGTTAGCGCCAC 4920  
Db 4861 TTCCCAATCAGCTAATGTTTAAAGCCCTTGATGATTTTCTCAGCGGGGTTAGCGCCAC 4920  
Qy 4921 GCAAAACAAGAAATGTGAAGCGGGAAGAGAAATGATCAGGATCGGGTAGGGATGG 4980  
Db 4921 GCAAAACAAGAAATGTGAAGCGGGAAGAGAAATGATCAGGATCGGGTAGGGATGG 4980  
Qy 4981 AGTGAATAAATCTTATCAAGAAATCTTTTGGGTAAATATCAAACTCAGGCAGGAACGA 5040  
Db 4981 AGTGAATAAATCTTATCAAGAAATCTTTTGGGTAAATATCAAACTCAGGCAGGAACGA 5040  
Qy 5041 GGAATACACATCATCGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCAACTCAAC 5100  
Db 5041 GGAATACACATCATCGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCAACTCAAC 5100  
Qy 5101 AGGCCATCTCTCAATCTTAAGTACTCAGCACTAGCTATTCTTGATGATTTCTTTTGGTT 5160  
Db 5101 AGGCCATCTCTCAATCTTAAGTACTCAGCACTAGCTATTCTTGATGATTTCTTTTGGTT 5160  
Qy 5161 GGGGTGTAATGTTGGTCTGTCAATGCGAGTTATGCTGAGGAAGATCAAACTCAGCT 5220  
Db 5161 GGGGTGTAATGTTGGTCTGTCAATGCGAGTTATGCTGAGGAAGATCAAACTCAGCT 5220  
Qy 5221 ATATCAAAAGGCTAATTTGAGGCTCAAGGAAAGGTATGTTTAAATCACTATCTTTTGG 5280  
Db 5221 ATATCAAAAGGCTAATTTGAGGCTCAAGGAAAGGTATGTTTAAATCACTATCTTTTGG 5280  
Qy 5281 ATCTTTTAAAGATCTTCAAGATGTCAGATGTTTATGTTGGCTTCTGGTGATTTTA 5340  
Db 5281 ATCTTTTAAAGATCTTCAAGATGTCAGATGTTTATGTTGGCTTCTGGTGATTTTA 5340  
Qy 5341 TGTTTTATAGAAATTTGGTCTTATATATTTGGCTATATATAGAGGTGGGTGATATGTA 5400  
Db 5341 TGTTTTATAGAAATTTGGTCTTATATATTTGGCTATATATAGAGGTGGGTGATATGTA 5400  
Qy 5401 GAATTCAGAGTTGATGTTGGAACCTTTTGTGCTTCAATTGAATAATCATCGAATCT 5460  
Db 5401 GAATTCAGAGTTGATGTTGGAACCTTTTGTGCTTCAATTGAATAATCATCGAATCT 5460  
Qy 5461 CAATTTCTTGGAGCCATATATGAGACATTTGAGACATCTATAGAACATATATGTAATGTA 5520  
Db 5461 CAATTTCTTGGAGCCATATATGAGACATTTGAGACATCTATAGAACATATATGTAATGTA 5520  
Qy 5521 TATTAACGTPACTTAAGTGAATTTTATGACAAAGTAAATAAATTTATGCCGAATGTACA 5580  
Db 5521 TATTAACGTPACTTAAGTGAATTTTATGACAAAGTAAATAAATTTATGCCGAATGTACA 5580  
Qy 5581 TGCTAATATCGAGTTTAAACTATTTTCCAAATATCAACTATTTTCTTTTCGTCCTCA 5640  
Db 5581 TGCTAATATCGAGTTTAAACTATTTTCCAAATATCAACTATTTTCTTTTCGTCCTCA 5640  
Qy 5641 CTTATATCTCTTATCTGATCTTATTTTCTTTTAAATTTCTTTTCTTTTCCCTTCCCAA 5700  
Db 5641 CTTATATACTCTTATCTGATCTTATTTTCTTTTAAATTTCTTTTCTTTTCCCTTCCCAA 5700  
Qy 5701 GACACAAAAAATAACAGAAACGAAAAAAGAGATTTTAAAAATTCATAACCCAC 5760  
Db 5701 GACACAAAAAATAACAGAAACGAAAAAAGAGATTTTAAAAATTCATAACCCAC 5760  
Qy 5761 GAGAAATATGCACTTAATTTCACTAATCCCAATTTTCAAGAAATTTATGATTTTGG 5820

Db 5761 GAGAAATATGCACTTAATTTCACTAATCCCAAAATTTTCAAGAAATTTATGATTTTGG 5820  
Qy 5821 CGATTTAATATGTTGTTTCACAATCATATGTCCTAACTAACTAAATTTGAAAAAGCAATGAA 5880  
Db 5821 CGATTTAATATGTTGTTTCACAATCATATGTCCTAACTAACTAAATTTGAAAAAGCAATGAA 5880  
Qy 5881 TGACTGAAACCAATGATCAATCTCTCAAGTCTCAACCTATGAAGAAATCATGTAACCAATAG 5940  
Db 5881 TGACTGAAACCAATGATCAATCTCTCAAGTCTCAACCTATGAAGAAATCATGTAACCAATAG 5940  
Qy 5941 ACTATCATCATGATTAAGTAAATGATATTAATGATTTCTTTTGAACATAGATATGTC 6000  
Db 5941 ACTATCATCATGATTAAGTAAATGATATTAATGATTTCTTTTGAACATAGATATGTC 6000  
Qy 6001 ATTTATCTGGATATTAAGATCGCGTTTAACTCTTTTGAAGAAATTTTCTGTTATATCTTTCT 6060  
Db 6001 ATTTATCTGGATATTAAGATCGCGTTTAACTCTTTTGAAGAAATTTTCTGTTATATCTTTCT 6060  
Qy 6061 TCTAATACATATGATCAATACACTTTTCTTTTAAAGAAATTTAAAACTTTATTTCAAAC 6120  
Db 6061 TCTAATACATATGATCAATACACTTTTCTTTTAAAGAAATTTAAAACTTTATTTCAAAC 6120  
Qy 6121 ATCGATCACTATTTTATCTTTTCTTGAATGACTACATTTATAGGCTCACACTTTGT 6180  
Db 6121 ATCGATCACTATTTTATCTTTTCTTGAATGACTACATTTATAGGCTCACACTTTGT 6180  
Qy 6181 TTCCGATCTAGATATCAATCAATCCACTTGTCTTTGAAGAGCGTGTGGAAGCTCTTCT 6240  
Db 6181 TTCCGATCTAGATATCAATCAATCCACTTGTCTTTGAAGAGCGTGTGGAAGCTCTTCT 6240  
Qy 6241 TTTTCCAGATGTTTCTCGTGGGTGGGGTCCACTCTTTTGGGACCACTGTCGGTAGAGGCAT 6300  
Db 6241 TTTTCCAGATGTTTCTCGTGGGTGGGGTCCACTCTTTTGGGACCACTGTCGGTAGAGGCAT 6300  
Qy 6301 CTTGAACGATAGCTTTCTTTTATCGCAATGATGCAATTTGTAAGGCCATCTTCTCTTTT 6360  
Db 6301 CTTGAACGATAGCTTTCTTTTATCGCAATGATGCAATTTGTAAGGCCATCTTCTCTTTT 6360  
Qy 6361 CTACTGCTCTTTCGATGAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCG 6420  
Db 6361 CTACTGCTCTTTCGATGAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCCG 6420  
Qy 6421 ATATTACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGA 6480  
Db 6421 ATATTACCTTTGTTGAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTTGA 6480  
Qy 6481 TATTCTTGGAGTAGACGAGAGTGTGCTGCCCACTGTTGGGGATCTAGATATCACATC 6540  
Db 6481 TATTCTTGGAGTAGACGAGAGTGTGCTGCCCACTGTTGGGGATCTAGATATCACATC 6540  
Qy 6541 AATCCACTTGTCTTGAAGCGTGGTTGGAACGCTCTTTTTCACGATGTTCTCGTGG 6600  
Db 6541 AATCCACTTGTCTTGAAGCGTGGTTGGAACGCTCTTTTTCACGATGTTCTCGTGG 6600  
Qy 6601 GTGGGGTCCACTTTTGGGACCACTGTGCGTAGAGGCATCTTGAACGATAGCTTTCTCT 6660  
Db 6601 GTGGGGTCCACTTTTGGGACCACTGTGCGTAGAGGCATCTTGAACGATAGCTTTCTCT 6660  
Qy 6661 TATCGCAATGATGGAATTTGTAAGAACCACTTCTCTTTTCTATCTGTCTTTCGATGAAGT 6720  
Db 6661 TATCGCAATGATGGAATTTGTAAGAACCACTTCTCTTTTCTATCTGTCTTTCGATGAAGT 6720  
Qy 6721 GACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATACCTTTGTTGAAAAG 6780  
Db 6721 GACAGATAGCTGGGCAATGGAATCCGAGGAGTTTCCGATATATACCTTTGTTGAAAAG 6780  
Qy 6781 TCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGAG 6840  
Db 6781 TCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGAG 6840  
Qy 6841 TGTGCTGCTCCACCAATGTTGGGATCTAGATATCAATCAATCCACTTCTGTTTGAAGACG 6900  
Db 6841 TGTGCTGCTCCACCAATGTTGGGATCTAGATATCAATCAATCCACTTCTGTTTGAAGACG 6900

6901 TGGTTGGAACGCTTCTTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGAC 6960  
6901 TGGTTGGAACGCTTCTTTTCCAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGAC 6960  
6961 CACTGTCGTAGAGGATCTTGAACGATAGCTTTCCCTTATCGCAATGATGCAATTTGT 7020  
6961 CACTGTCGTAGAGGATCTTGAACGATAGCTTTCCCTTATCGCAATGATGCAATTTGT 7020  
7021 AGAAGCAATCTTCTTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGA 7080  
7021 AGAAGCAATCTTCTTTTCTACTGTCCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGA 7080  
7081 ATCCGAGGAGGTTTCCCGATATTAACCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTT 7140  
7081 ATCCGAGGAGGTTTCCCGATATTAACCTTTGTTGAAAAGTCTCAATAGCCCTCTGGTCTTT 7140  
7141 CTGAGACTGATCTTTGATATTTCTGGAGTACGAGAGTGCTGCTCCACCATGTTGG 7200  
7141 CTGAGACTGATCTTTGATATTTCTGGAGTACGAGAGTGCTGCTCCACCATGTTGG 7200  
7201 GGATCTAGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGAACTCTTCTTTT 7260  
7201 GGATCTAGATATCACATCAATCCACTTGTCTTGAAGACGTGGTGGAACTCTTCTTTT 7260  
7261 CCACGATGTTCTCGTGGGTGGGGTCCATCTTTTGGGACCACTGTCGGTAGAGGCATCTT 7320  
7261 CCACGATGTTCTCGTGGGTGGGGTCCATCTTTTGGGACCACTGTCGGTAGAGGCATCTT 7320  
7321 GAACGATAGCTTCTTATCGCAATGATGCGATTTGAGAGCCATCTTCTTTCTA 7380  
7321 GAACGATAGCTTCTTATCGCAATGATGCGATTTGAGAGCCATCTTCTTTCTA 7380  
7381 CTGTCTTTTGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATA 7440  
7381 CTGTCTTTTGAAGTACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCGATA 7440  
7441 TTACCTTTTGAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATAT 7500  
7441 TTACCTTTTGAAGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATAT 7500  
7501 TCTTGAGTAGAGGAGTGCTGCTCCACCATGTTGGGATCCACTAGTCTAGAGCG 7560  
7501 TCTTGAGTAGAGGAGTGCTGCTCCACCATGTTGGGATCCACTAGTCTAGAGCG 7560  
7561 GCGCCACCGCGTGAGCT 7580  
7561 GCGCCACCGCGTGAGCT 7580

RESULT 2  
US-10-650-249-18  
; Sequence 18, Application US/10650249  
; Publication No. US2004045055A1  
; GENERAL INFORMATION:  
; APPLICANT: Neff, Michael M  
; TITLE OF INVENTION: THE GENE FOR A DOF TRANSCRIPTION FACTOR CAPABLE OF ALTERING  
; TITLE OF INVENTION: THE SIZE AND STATURE OF A PLANT  
; FILE REFERENCE: WSHU 2064.1  
; CURRENT APPLICATION NUMBER: US/10/650,249  
; CURRENT FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US 60/406,657  
; PRIOR FILING DATE: 2002-08-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 1746  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-650-249-18

Query Match 19.8%; Score 1503; DB 7; Length 1746;  
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1743; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
QY 4436 AACAGAGAGAAATCCAGATCGAAATCTACGGTCTGGTCTCGACTGATTAATCTACT 4495  
DB 1 AACAGAGAGAAATCCAGATCGAAATCTACGGTCTGGTCTCGACTGATTAATCTACT 60  
QY 4496 AGTACTTTCATCTTACTTCTCGCCCAAGTTACTCAAAACCTTACAAAGTTTTCATAGCTAC 4555  
DB 61 AGTACTTTCATCTTACTTCTCGCCCAAGTTACTCAAAACCTTACAAAGTTTTCATAGCTAC 120  
QY 4556 GGTCAAAATCCCGAGTTTAAATTCCAAATCTTGGCTCTCTCCAAAGCCCTTGA 4615  
DB 121 GGTCAAAATCCCGAGTTTAAATTCCAAATCTTGGCTCTCTCCAAAGCCCTTGA 180  
QY 4616 GATTAACAATCAAGCAACACATCGATTTAGTTTGGTGGAACTCAAAATAGCAACATGATA 4675  
DB 181 GATTAACAATCAAGCAACACATCGATTTAGTTTGGTGGAACTCAAAATAGCAACATGATA 240  
QY 4676 AGTGTATGAGTTCTAGTGGTGGGATCTTGGATGATGGAGAAATACCTCCATCAACAA 4735  
DB 241 AGTGTATGAGTTCTAGTGGTGGGATCTTGGATGATGGAGAAATACCTCCATCAACAA 300  
QY 4736 GCTCAGCAATCCCTTCTTGTATCAACACTACCGGATTTGGTGGAACTCTTCAACCGGTTA 4795  
DB 301 GCTCAGCAATCCCTTCTTGTATCAACACTACCGGATTTGGTGGAACTCTTCAACCGGTTA 360  
QY 4796 TATCAATTACTAGAAAGTAAAGGAGGTGTTAATCAAGGTGATTTCAACAGAGAGTAGT 4855  
DB 361 TATCAATTACTAGAAAGTAAAGGAGGTGTTAATCAAGGTGATTTCAACAGAGAGTAGT 420  
QY 4856 GATTAATCCATCAGCTAATGTTTAAAGCCCTTGGTGGATTTTCTTCAAGCGGGTTAGC 4915  
DB 421 GATTAATCCATCAGCTAATGTTTAAAGCCCTTGGTGGATTTTCTTCAAGCGGGTTAGC 480  
QY 4916 GCCAGCAAAACAAAGAAATGTTGAAGCGGAGAGAAATGATCAGGATCGGGTAGGATGG 4975  
DB 481 GCCAGCAAAACAAAGAAATGTTGAAGCGGAGAGAAATGATCAGGATCGGGTAGGATGG 540  
QY 4976 GATGAGTGAATTAATCAAGAAATCTTTTGGGTAAATCAACATAAACTCAGGACAG 5035  
DB 541 GATGAGTGAATTAATCAAGAAATCTTTTGGGTAAATCAACATAAACTCAGGACAG 600  
QY 5036 AACGAGGATATACATATCATGCGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCCAAC 5095  
DB 601 AACGAGGATATACATATCATGCGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCCAAC 660  
QY 5096 TCAACAGGCGCATCTCTCATTTC--TAAAGTACTCAGCACTAGCTATTTCTGATGATTTCTTT 5152  
DB 661 TCAACAGGCGCATCTCTCATTTC--TAAAGTACTCAGCACTAGCTATTTCTGATGATTTCTTT 720  
QY 5153 TGTGTTGGGTGTACATTTGGTCTTGTTCATGCGAGTTAATTTGTGAGGAAGATCAAAACC 5212  
DB 721 TGTGTTGGGTGTACATTTGGTCTTGTTCATGCGAGTTAATTTGTGAGGAAGATCAAAACC 780  
QY 5213 ATGCACTATATCCAAAGCTTAATTTTGGGCTTCAAGGAAAGGATGATGCTTATAAACA 5272  
DB 781 ATGCACTATATCCAAAGCTTAATTTTGGGCTTCAAGGAAAGGATGATGCTTATAAACA 840  
QY 5273 TCTTTTGTATCTTTTAAAGATCTTCAAAAGTGTGAGTATGTTTATTTGGTCTTCTGCT 5332  
DB 841 TCTTTTGTATCTTTTAAAGATCTTCAAAAGTGTGAGTATGTTTATTTGGTCTTCTGCT 900  
QY 5333 GATATTTATTTTATTTAGAAATTTGGTCTTATATATTTGGCTATATATAGAGGTGCGGTG 5392  
DB 901 GATATTTATTTTATTTAGAAATTTGGTCTTATATATTTGGCTATATATAGAGGTGCGGTG 960  
QY 5393 ATATGATGAATTCAGAGTTGATGTTGGAACTTTTGTGTTGTTTGTGTTTGTGTTTGTATCAT 5452  
DB 961 ATATGATGAATTCAGAGTTGATGTTGGAACTTTTGTGTTGTTTGTGTTTGTGTTTGTATCAT 1020  
QY 5453 CGAATTTCAATTTCTTGGAGACCCATATATGAGACATTTGAGACATCTATAGAACATATAT 5512  
DB 1021 CGAATTTCAATTTCTTGGAGACCCATATATGAGACATTTGAGACATCTATAGAACATATAT 1080



QY	5513	GTAATGTATATTAAGTACTTAAGTCGAATTTTATGACCAAGTAATAAATTAATGCG	5572
DB	1081	GTAATGTATATTAAGTACTTAAGTCGAATTTTATGACCAAGTAATAAATTAATGCG	1140
QY	5573	AATGTACATGCTAAATCGAGTTTAAACTATTTTCCAAATAACAACATATTTCTCT	5632
DB	1141	AATGTACATGCTAAATCGAGTTTAAACTATTTTCCAAATAACAACATATTTCTCT	1200
QY	5633	TCGTCCAACTTATATCTTATCTGATTTCTTAATTTCTTTTAAATTCCTTTTCC	5692
DB	1201	TCGTCCAACTTATATCTTATCTGATTTCTTAATTTCTTTTAAATTCCTTTTCC	1260
QY	5693	TTTCCCAAGACACAAAAAATAATACAGAAAAGAAAAAGAGATTTTAAAAATTCA	5752
DB	1261	TTTCCCAAGACACAAAAAATAATACAGAAAAGAAAAAGAGATTTTAAAAATTCA	1320
QY	5753	TAAACACAGAAATATGACCTTAAATTCAGACTAATCCCAAAATTTTCAAGAAATTA	5812
DB	1321	TAAACACAGAAATATGACCTTAAATTCAGACTAATCCCAAAATTTTCAAGAAATTA	1380
QY	5813	TATTTTGGGATTTAATATTTGTTTCACAATCAATCAATGAGCCCAACTAACTAAT	5872
DB	1381	TATTTTGGGATTTAATATTTGTTTCACAATCAATCAATGAGCCCAACTAATTA	1440
QY	5873	CAATGGAATGACTGAACCAATGATTAATCTCAAGTCTCAACCTATGAAGAAATCA	5932
DB	1441	CAATGGAATGACTGAACCAATGATTAATCTCAAGTCTCAACCTATGAAGAAATCA	1500
QY	5933	ACCAATAGACTCATGATGATTAATGATGATGATGATGATGATGATGATGATGAT	5992
DB	1501	ACCAATAGACTCATGATGATTAATGATGATGATGATGATGATGATGATGATGAT	1560
QY	5993	GATATGTCAATTTATCTGGATATAAAGATGGCGTTTAACTACTTTTGCAATTTTGT	6052
DB	1561	GATATGTCAATTTATCTGGATATAAAGATGGCGTTTAACTACTTTTGCAATTTTGT	1620
QY	6053	ATCTTTCTTCAATACATATGATCAATACATCTTTTGTGTTTAAAGAAATTAACCT	6112
DB	1621	ATCTTTCTTCAATACATATGATCAATACATCTTTTGTGTTTAAAGAAATTAACCT	1680
QY	6113	TTTCAACATCGATCAATTTTACTTTTGTGTTTCCATATGATGATGATGATGATG	6172
DB	1681	TTTCAACATCGATCAATTTTACTTTTGTGTTTCCATATGATGATGATGATGATG	1740
QY	6173	ACTTTT 6178	
DB	1741	ACTTTT 1746	

RESULT 3  
US-10-033-190-3/c  
; Sequence 3, Application US/10033190  
; Publication No. US20020133848A1  
; GENERAL INFORMATION:  
; APPLICANT: Exelixis Plant Sciences, Inc.  
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF AN ANTHOCYANIN MUTANT (ANT  
; FILE OF INVENTION: TOMATO  
; FILE REFERENCE: EP01-002C  
; CURRENT APPLICATION NUMBER: US/10/033,190  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: US 60/244,685  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 10078  
; TYPE: DNA  
; ORGANISM: pAG3202  
US-10-033-190-3

Query Match 18.5%; Score 1402; DB 5; Length 10078;  
Best Local Similarity 100.0%; Pred. No. 0;

		Matches 1402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	6179	GTTCGGATCTAGATATCACATCAATCCATGCTTTTGAAGACGCTGTTGGAACGCTTTC	6238
DB	4249	GTTCGGATCTAGATATCACATCAATCCATGCTTTTGAAGACGCTGTTGGAACGCTTTC	4190
QY	6239	TTTTTCCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTTGGGACCACTGTCTGGTAGAGC	6298
DB	4189	TTTTTCCACGATGTTCTCTGTTGGGTGGGGTCCATCTTTTGGGACCACTGTCTGGTAGAGC	4130
QY	6299	ATCTTGAACGATAGCTTTCTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTCT	6358
DB	4129	ATCTTGAACGATAGCTTTCTTTATCGCAATGATGGCAATTTGTAGAGCCATCTTCTCT	4070
QY	6359	TTCTACTGCTCTTTTCGATGAAGTGCAGATAGCTGGGCAATGGAATCCGAGAGGTTTCC	6418
DB	4069	TTCTACTGCTCTTTTCGATGAAGTGCAGATAGCTGGGCAATGGAATCCGAGAGGTTTCC	4010
QY	6419	CGATATTTACCTTTCTGAAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTT	6478
DB	4009	CGATATTTACCTTTCTGAAAAAGTCTCAATAGCCCTCTGGTCTTCTGAGACTGTATCTTT	3950
QY	6479	GATATTTTGGAGTAGACGAGAGTCTGCTGCCATGCTCCACATGTTGGGATCTAGATACAC	6538
DB	3949	GATATTTTGGAGTAGACGAGAGTCTGCTGCCATGCTCCACATGTTGGGATCTAGATACAC	3890
QY	6539	TCATATCCACTTCTTTGAAAGCTGTTGGAACGCTCTTCTTTTCCACGATGTTCTCGT	6598
DB	3889	TCATATCCACTTCTTTGAAAGCTGTTGGAACGCTCTTCTTTTCCACGATGTTCTCGT	3830
QY	6599	GGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCACTCTGAAACGATAGCTTTCC	6658
DB	3829	GGGTGGGGTCCATCTTTGGGACCACTGTCGGTAGAGGCACTCTGAAACGATAGCTTTCC	3770
QY	6659	TTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTCTACTGCTTTTCGATGAA	6718
DB	3769	TTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTCTACTGCTTTTCGATGAA	3710
QY	6719	GTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTAACCTTTGTTGAA	6778
DB	3709	GTGACAGATAGCTGGGCAATGGAATCCGAGAGGTTTCCCGATATTAACCTTTGTTGAA	3650
QY	6779	AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTGTATTTCTTGGAGTAGACGAG	6838
DB	3649	AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTGTATTTCTTGGAGTAGACGAG	3590
QY	6839	AGTGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCACTTTGTTGAAAG	6898
DB	3589	AGTGTGCTGCTCCACCATGTTGGGATCTAGATATCAATCAATCACTTTGTTGAAAG	3530
QY	6899	CGTGTGGAACGCTCTTCTTTTCCAGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGG	6958
DB	3529	CGTGTGGAACGCTCTTCTTTTCCAGATGTTCTCTGTTGGGTGGGGTCCATCTTTGGG	3470
QY	6959	ACCACTGCTGGTAGAGGCACTTTGAACGATAGCTTTCTTTTCCCAATGATGGCATTT	7018
DB	3469	ACCACTGCTGGTAGAGGCACTTTGAACGATAGCTTTCTTTTCCCAATGATGGCATTT	3410
QY	7019	GTAGAAGCCATCTTCTTTTCTACTGTCTCTTTCGATGAAGTGCAGATAGCTGGGCAATG	7078
DB	3409	GTAGAAGCCATCTTCTTTTCTACTGTCTCTTTCGATGAAGTGCAGATAGCTGGGCAATG	3350
QY	7079	GAAATCGAGAGGTTTCCCGATATTTACCCCTTTGTTGAAAGTCTCAATAGCCCTCTGTC	7138
DB	3349	GAAATCGAGAGGTTTCCCGATATTTACCCCTTTGTTGAAAGTCTCAATAGCCCTCTGTC	3290
QY	7139	TTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTT	7198
DB	3289	TTCTGAGACTGTATCTTTGATATTTCTTGGAGTAGACGAGAGTGTCTGCTCCACCATGTT	3230
QY	7199	GGGGATCTAGATATCACATCAATCCACTTTGTTGAAGACGTTGGTTGGAACGCTTCTCT	7258
DB	3229	GGGGATCTAGATATCACATCAATCCACTTTGTTGAAGACGTTGGTTGGAACGCTTCTCT	3170



QY 7259 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGCATC 7318  
DB 3169 TTCCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGCATC 3110  
QY 7319 TTGAACGATAGCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 7378  
DB 3109 TTGAACGATAGCTTTCTTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 3050  
QY 7379 TACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGA 7438  
DB 3049 TACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGA 2990  
QY 7439 TATTACCTTTGTGTAAGTCTCAATAGCCCTCTCGTCTTTCTGAGACTGTATCTTTGAT 7498  
DB 2989 TATTACCTTTGTGTAAGTCTCAATAGCCCTCTCGTCTTTCTGAGACTGTATCTTTGAT 2930  
QY 7499 ATTCTTGGAGTACACGAGAGTGTGCTCCACCATGTTGGGATCCACTAGTTCTAGAG 7558  
DB 2929 ATTCTTGGAGTACACGAGAGTGTGCTCCACCATGTTGGGATCCACTAGTTCTAGAG 2870  
QY 7559 CGGCCGCCACCGCGGTGGAGCT 7580  
DB 2869 CGGCCGCCACCGCGGTGGAGCT 2848

## RESULT 4

US-09-522-334-1/c  
; Sequence 1, Application US/09522334  
; Patent No. US20020157130A1  
; GENERAL INFORMATION:  
; APPLICANT: Wagner, Ry  
; APPLICANT: Mathews, Helena  
; APPLICANT: Liu, Xing liang  
; APPLICANT: Waggoner, Wency J.  
; TITLE OF INVENTION: TRAIT-ASSOCIATED GENE IDENTIFICATION  
; TITLE OF INVENTION: METHOD  
; FILE REFERENCE: 4257-0018.30  
; CURRENT APPLICATION NUMBER: US/09/522,334  
; CURRENT FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: US 60/124,232  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1361  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: modified enhancer  
US-09-522-334-1

Query Match 17.9%; Score 1360; DB 3; Length 1361;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6179 GTTTCGGATCTAGATATCACATCAATCCACTTGTGTAAGACGTTGGTGAACGTCCTTC 6238  
DB 1360 GTTTCGGATCTAGATATCACATCAATCCACTTGTGTAAGACGTTGGTGAACGTCCTTC 1301  
QY 6239 TTTTTCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGC 6298  
DB 1300 TTTTTCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGC 1241  
QY 6299 ATCTTGAACGATAGCTTCTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTT 6358  
DB 1240 ATCTTGAACGATAGCTTCTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTT 1181  
QY 6359 TTCTACTGTCTTTGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCC 6418  
DB 1180 TTCTACTGTCTTTGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCC 1121  
QY 6419 CGATATTACCTTTGTGTAAGTCTCAATAGCCCTCTCGTCTTCTGAGACTGTATCTTTT 6478

DB 1120 CGATATTACCTTTGTGTAAGTCTCAATAGCCCTCTCGTCTTCTGAGACTGTATCTTTT 1061  
QY 6479 GATATTCTTTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGGATCTAGATATCACA 6538  
DB 1060 GATATTCTTTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTTGGGGATCTAGATATCACA 1001  
QY 6539 TCAATCCACTTGTGTAAGACGTTGGTGGAAAGTGTGCTTCTTTTCCAGATGTCCTCGT 6598  
DB 1000 TCAATCCACTTGTGTAAGACGTTGGTGGAAAGTGTGCTTCTTTTCCAGATGTCCTCGT 941  
QY 6599 GGTGGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGCATCTTGAAACGATAGCCCTTTC 6658  
DB 940 GGTGGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGCATCTTGAAACGATAGCCCTTTC 881  
QY 6659 TTTATCGCAATGATGGCATTTGTAGAACCATCTTCTTCTTCTACTGTCTTTCGATGAA 6718  
DB 880 TTTATCGCAATGATGGCATTTGTAGAACCATCTTCTTCTTCTACTGTCTTTCGATGAA 821  
QY 6719 GTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATTTACCTTTTGTGAAA 6778  
DB 820 GTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGATATTTACCTTTTGTGAAA 761  
QY 6779 AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTTCTCGAGTAGACGAG 6838  
DB 760 AGTCTCAATAGCCCTCTGCTCTTCTGAGACTGTATCTTTGATATTTCTTGAGTAGACGAG 701  
QY 6839 AGTCTGCTGCTCCACCATGTTGGGATCTAGATATCACAATCAATCCACTTGTCTTGAAGA 6898  
DB 700 AGTCTGCTGCTCCACCATGTTGGGATCTAGATATCACAATCAATCCACTTGTCTTGAAGA 641  
QY 6899 CTGCTTGGAACTCTTCTTTTTCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGG 6958  
DB 640 CTGCTTGGAACTCTTCTTTTTCACGATGTTCTCGTGGGTGGGGTCCATCTTTGGG 581  
QY 6959 ACCACTGCTGGTAGAGGCATCTTGAAACGATAGCTTCTTCTTATCGCAATGAGGCATTT 7018  
DB 580 ACCACTGCTGGTAGAGGCATCTTGAAACGATAGCTTCTTCTTATCGCAATGAGGCATTT 521  
QY 7019 GTAGAAGCATCTTCTTTTCTGATGTTCTTCTGATGAAGTGCAGATAGCTGGGCAATG 7078  
DB 520 GTAGAAGCATCTTCTTTTCTGATGTTCTTCTGATGAAGTGCAGATAGCTGGGCAATG 461  
QY 7079 GAATCCGAGGAGGTTTCCGATATTAACCTTTGTGAAAGTCTCAATAGCCCTCTGCTC 7138  
DB 460 GAATCCGAGGAGGTTTCCGATATTAACCTTTGTGAAAGTCTCAATAGCCCTCTGCTC 401  
QY 7139 TTCTGAGACTGTATCTTTGATATTTCTTGAGTAGACGAGTGTGCTGCTCCACCATGTT 7198  
DB 400 TTCTGAGACTGTATCTTTGATATTTCTTGAGTAGACGAGTGTGCTGCTCCACCATGTT 341  
QY 7199 GGGGATCTAGATATCACAATCAATCCACTTGTGTAAGACGTTGGAACTGTTCTTCTT 7258  
DB 340 GGGGATCTAGATATCACAATCAATCCACTTGTGTAAGACGTTGGAACTGTTCTTCTT 281  
QY 7259 TTCCAAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGCATC 7318  
DB 280 TTCCAAGATGTTCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTGCGGTAGAGGCATC 221  
QY 7319 TTGAACGATAGCTTCTTCTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 7378  
DB 220 TTGAACGATAGCTTCTTCTTATCGCAATGATGGCAATTTGTAGAAGCCATCTTCTTTTC 161  
QY 7379 TACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGA 7438  
DB 160 TACTGTCTTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGA 101  
QY 7439 TATTACCTTTGTGTAAGTCTCAATAGCCCTCTGCTTCTGAGACTGTATCTTTGAT 7498  
DB 100 TATTACCTTTGTGTAAGTCTCAATAGCCCTCTGCTTCTGAGACTGTATCTTTGAT 41  
QY 7499 ATTCTTGGAGTAGACGAGAGTGTGCTGCTCCACCATGTTG 7538



QY 5237 TTTGAGGCTCAAGGAAGGTATGGTTATATAAACTATCTTTTGATCTTTTAAAGATCT 5296  
Db 121 TTTGAGGCTCAAGGAAGGTATGGTTATATAAACTATCTTTTGATCTTTTAAAGATCT 180  
QY 5297 TCAAAGTGTGAGTATGTTATGTTGCTTCTGGTGATATTTATGTTTATTTAGAAATTT 5356  
Db 181 TCAAAGTGTGAGTATGTTATGTTGCTTCTGGTGATATTTATGTTTATTTAGAAATTT 240  
QY 5357 GGTCTTATATATTGGCTATATATAGAGGTTGGTGATATGATGAATTCAGAGTTGAT 5416  
Db 241 GGTCTTATATATTGGCTATATATAGAGGTTGGTGATATGATGAATTCAGAGTTGAT 300  
QY 5417 GTTGGAAACTTTTGTGCTTCAATTAATCAATCAATTCATCAATTTCTTGGAGACC 5476  
Db 301 GTTGGAAACTTTTGTGCTTCAATTAATCAATCAATTCATCAATTTCTTGGAGACC 360  
QY 5477 CATTTATGAGACATTTGAGACATCTATAGAACATATATGTAATGATATTAACCTACTTAA 5536  
Db 361 CATTTATGAGACATTTGAGACATCTATAGAACATATATGTAATGATATTAACCTACTTAA 420  
QY 5537 GTCGAATTTTATGACCAAGTAATTAATATGCGGAATGTAATGCTAATATCGAGTTT 5596  
Db 421 GTCGAATTTTATGACCAAGTAATTAATATGCGGAATGTAATGCTAATATCGAGTTT 480  
QY 5597 AACTATTTTTCGAATATACAACTATTTCTTCTTTCGTCGAATTTATATCTTAT 5656  
Db 481 AACTATTTTTCGAATATACAACTATTTCTTCTTTCGTCGAATTTATATCTTAT 540  
QY 5657 CTGATTTCTTATTTCTTCTTTTAAATTCCTTTTCCCAAGACACAAAAA 5716  
Db 541 CTGATTTCTTATTTCTTCTTTTAAATTCCTTTTCCCAAGACACAAAAA 600  
QY 5717 AATACAGAAACGAAAAAGAGATTTTAAATAATTCATAACCCACGAAATTTACACCTA 5776  
Db 601 AATACAGAAACGAAAAAGAGATTTTAAATAATTCATAACCCACGAAATTTACACCTA 660  
QY 5777 AATTGAGACTAATCCCGCAATTTTCAAGAAATTTATGATATTTTGGCAATTTATTTGT 5836  
Db 661 AATTGAGACTAATCCCGCAATTTTCAAGAAATTTATGATATTTTGGCAATTTATTTGT 720  
QY 5837 TCACAAATCAATAGGCACTAATTAATGAAAGCAATGGAATGCAATGAAACCATGCA 5896  
Db 721 TCACAAATCAATAGGCACTAATTAATGAAAGCAATGGAATGCAATGAAACCATGCA 780  
QY 5897 TAATCTCTCAAGTCTCAACCTATGAAGAAATCATGTAAACCAATGACTATCATGATTA 5956  
Db 781 TAATCTCTCAAGTCTCAACCTATGAAGAAATCATGTAAACCAATGACTATCATGATTA 840  
QY 5957 GTTAATGCAATGATCTAATGATATCTTTGAAACATAGATATGTCATTTATCTGGATATA 6016  
Db 841 GTTAATGCAATGATCTAATGATATCTTTGAAACATAGATATGTCATTTATCTGGATATA 900  
QY 6017 AGATGGGTTTAACTTACTTTCGAATTTTGTATATCTTCTTCTTAATACATATGATC 6076  
Db 901 AGATGGGTTTAACTTACTTTCGAATTTTGTATATCTTCTTCTTAATACATATGATC 960  
QY 6077 AATACACTTTTGTGTTTAAAAAGAAATTAATACTTATTTCAACATCGATCAATTTTAA 6136  
Db 961 AATACACTTTTGTGTTTAAAAAGAAATTAATACTTATTTCAACATCGATCAATTTTAA 1020  
QY 6137 CTTTGTGTTCCATATGACTACATTTTATAGGCTCACACTTTT 6178  
Db 1021 CTTTGTGTTCCATATGACTACATTTTATAGGCTCACACTTTT 1062

## RESULT 7

US-10-650-249-19  
; Sequence 19, Application US/10650249  
; Publication No. US20040045055A1  
; GENERAL INFORMATION:  
; APPLICANT: Neff, Michael M  
; TITLE OF INVENTION: THE GENE FOR A DOF TRANSCRIPTION FACTOR CAPABLE OF ALTERING  
; TITLE OF INVENTION: THE SIZE AND STATURE OF A PLANT

FILE REFERENCE: WSHU 2064.1  
; CURRENT APPLICATION NUMBER: US/10/650,249  
; PRIOR FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US 60/406,657  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 1058  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-650-249-19

Query Match 14.0%; Score 1058; DB 7; Length 1058;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1058; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3045 CAACCAAGAACGATGACGTATATGATTGACTTGCACAAAATAAGCAAAACAAATACCTGTT 3104  
Db 1 CAACCAAGAACGATGACGTATATGATTGACTTGCACAAAATAAGCAAAACAAATACCTGTT 60  
QY 3105 CAAATCGACACTTAATTTCCAAAAAGGTTAGTAATAGTAAGAGGCTTTTATTATGAAA 3164  
Db 61 CAAATCGACACTTAATTTCCAAAAAGGTTAGTAATAGTAAGAGGCTTTTATTATGAAA 120  
QY 3165 ACAAAAGAAATTAAGAGGCTTAAGAGATGATGAAATTTGAAAGAGAAAAAGAGCAATTG 3224  
Db 121 ACAAAAGAAATTAAGAGGCTTAAGAGATGATGAAATTTGAAAGAGAAAAAGAGCAATTG 180  
QY 3225 TTATAGAAAGAAAAAGAGAGAGTAAGAGATTAAGAGAACACAAATAAATTAACCA 3284  
Db 181 TTATAGAAAGAAAAAGAGAGAGTAAGAGATTAAGAGAACACAAATAAATTAACCA 240  
QY 3285 AGGAAACTTCATTTCTTTTATCCCAATTCAGCTCTCCCTTCTCTCTCTCTCTCTCTCT 3344  
Db 241 AGGAAACTTCATTTCTTTTATCCCAATTCAGCTCTCCCTTCTCTCTCTCTCTCTCTCT 300  
QY 3345 CTCTCTCTCTAGATCAATTTCTTTCTATGATGATATATCCATATCTGCGACC 3404  
Db 301 CTCTCTCTCTAGATCAATTTCTTTCTATGATGATATATCCATATCTGCGACC 360  
QY 3405 TCTTACCTTAAAGGATCAAGTAAAGATTCAGAGATTCAGAGATGTTTCTCATCTCTTCCAGTG 3464  
Db 361 TCTTACCTTAAAGGATCAAGTAAAGATTCAGAGATGTTTCTCATCTCTTCCAGTG 420  
QY 3465 AATCAGTTTCAATTTCCCAAAATTTGGCAGAGTAAATAATCAGTTTATGATATTTGCTAGAT 3524  
Db 421 AATCAGTTTCAATTTCCCAAAATTTGGCAGAGTAAATAATCAGTTTATGATATTTGCTAGAT 480  
QY 3525 GTTCTGATTCGTTCTTTTCTCTCAAGCTCGATCAAGATTTATGAAAATTTGATGAGA 3584  
Db 481 GTTCTGATTCGTTCTTTTCTCTCAAGCTCGATCAAGATTTATGAAAATTTGATGAGA 540  
QY 3585 TTTTGTTCGACAAAATTCCTAGCTATTTGGAGCGGCATATATATTTACTTATGAAATTTTC 3644  
Db 541 TTTTGTTCGACAAAATTCCTAGCTATTTGGAGCGGCATATATATTTACTTATGAAATTTTC 600  
QY 3645 TTAGTTGATTAACCCCTTTTCTTCTCTCGAATATACGAAAATATATAAAGAT 3704  
Db 601 TTAGTTGATTAACCCCTTTTCTTCTCTCGAATATACGAAAATATATAAAGAT 660  
QY 3705 GATTTCAATTTTGGTCTTTTCTTCTCAAGACTTTTAAAAAATTTTCTTAGTTGA 3764  
Db 661 GATTTCAATTTTGGTCTTTTCTTCTCAAGACTTTTAAAAAATTTTCTTAGTTGA 720  
QY 3765 TAAAAACCTTTTCTTGTCTTCTCCAGGCTTATGATATATTTTCTTACAGGAT 3824  
Db 721 TAAAAACCTTTTCTTGTCTTCTCCAGGCTTATGATATATTTTCTTACAGGAT 780  
QY 3825 AATTTTCTTTGTTAGATTTTACACCGCATGAAATTTATCATTCAAAAAATAAAAA 3884  
Db 781 AATTTTCTTTGTTAGATTTTACACCGCATGAAATTTATCATTCAAAAAATAAAAA 840



Db 241 AGTGGTATGAGTTCTAGTGGTGGATCTTGGATGCATGGAGATACCTCCATCACAAAC 300  
QY 4736 GCTCAGCAATCCCTTTCTTGATCAACATACCGGATTTGGTGAATCTTCAACCGGTGA 4795  
Db 301 GCTCAGCAATCCCTTTCTTGATCAACATACCGGATTTGGTGAATCTTCAACCGGTGA 360  
QY 4796 TATCCATTACTAAGAGTAAAGGAGGTGTAAATCAAGGTGATTTCTCAACAGAGAGTAGT 4855  
Db 361 TATCCATTACTAAGAGTAAAGGAGGTGTAAATCAAGGTGATTTCTCAACAGAGAGTAGT 420  
QY 4856 GATTATTTCCAAATCAGTGAATTTTAAAGCCCTTGATGGATTTTCTTCAAGCGGGTTAGC 4915  
Db 421 GATTATTTCCAAATCAGTGAATTTTAAAGCCCTTGATGGATTTTCTTCAAGCGGGTTAGC 480  
QY 4916 GCCACGAAACAAAGAAATGTAAAGCGGAGAGAAATGATCAGGATCGGGGTAGGGATGG 4975  
Db 481 GCCACGAAACAAAGAAATGTAAAGCGGAGAGAAATGATCAGGATCGGGGTAGGGATGG 540  
QY 4976 GATGGAGTGAATTAATCAAGAACTTTTGGGTAAATTAACATAACTCAGGCAGG 5035  
Db 541 GATGGAGTGAATTAATCAAGAACTTTTGGGTAAATTAACATAACTCAGGCAGG 600  
QY 5036 AACGAGGAATACATCATCGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCAAAC 5095  
Db 601 AACGAGGAATACATCATCGGGAGGTAAACAGTTCTTGGACCGGTTTCACTCCAAAC 660  
QY 5096 TCAACAGGCCATCTCTCATTTAA 5119  
Db 661 TCAACAGGCCATCTCTCATTTAA 684

## RESULT 10

US-10-487-901-7187  
; Sequence 7187, Application US/10487901  
; Publication No. US20050091708A1  
; GENERAL INFORMATION:  
; APPLICANT: Oreido, Jeremiah Vincent  
; APPLICANT: McCrery, David  
; APPLICANT: Pell, Randy  
; APPLICANT: Miller, Barbara  
; APPLICANT: Weglarz, Thaddeus  
; APPLICANT: Gachotte, Daniel  
; APPLICANT: Blakeslee, Beth  
; APPLICANT: Reddy, Avutu  
; APPLICANT: Shukla, Vipula  
; APPLICANT: Crosley, Rodney  
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri  
; FILE REFERENCE: DOW-08552  
; CURRENT APPLICATION NUMBER: US/10/487,901  
; CURRENT FILING DATE: 2004-02-26  
; NUMBER OF SEQ ID NOS: 7560  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7187  
; LENGTH: 645  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-487-901-7187

Query Match 7.7%; Score 581; DB 9; Length 645;

Best Local Similarity 100.0%; Pred. No. 5.7e-252;

Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4232 ATGGTGAACGTCTCGGATCGGAAAGTCCGATTCGCTGAAGCAGCTCTAAATGCCCT 4291  
Db 6 ATGGTGAACGTCTCGGATCGGAAAGTCCGATTCGCTGAAGCAGCTCTAAATGCCCT 65

QY 4292 AGATGTGACTCAACCACTAAGTTCTGTACTTCAATAACTATAGCCTTACTCAACCT 4351  
Db 66 AGATGTGACTCAACCACTAAGTTCTGTACTTCAATAACTATAGCCTTACTCAACCT 125

QY 4352 CGCCATTTCTGCARACATGTCGTGCTATTGGACACGTGGCGGTTCCCTTGAGGAATGTT 4411  
Db 126 CGCCATTTCTGCARACATGTCGTGCTATTGGACACGTGGCGGTTCCCTTGAGGAATGTT 185  
QY 4412 CTGTGTGAGGAGGCTTTAGGAGGAAACAAGAGAAACAAATCCAGATCGAAATCTACGGTC 4471  
Db 186 CTGTGTGAGGAGGCTTTAGGAGGAAACAAGAGAAACAAATCCAGATCGAAATCTACGGTC 245  
QY 4472 GTGGTCTCGACTGATTAATCTACTAGTAATTCAATCACTTCTTCTCGCCCAAGTTACTCA 4531  
Db 246 GTGGTCTCGACTGATTAATCTACTAGTAATTCAATCACTTCTTCTCGCCCAAGTTACTCA 305  
QY 4532 AACCTTAGCAAGTTTCATAGCTACGTCAAATCCCGGAGTTTAAATCCCACTTGGCCATC 4591  
Db 306 AACCTTAGCAAGTTTCATAGCTACGTCAAATCCCGGAGTTTAAATCCCACTTGGCCATC 365  
QY 4592 TTGGCTCTCTCTCAAGCCCTTGGAGATTACAATTTCAAGCAACACTGGATTTTGGT 4651  
Db 366 TTGGCTCTCTCTCAAGCCCTTGGAGATTACAATTTCAAGCAACACTGGATTTTGGT 425  
QY 4652 GGAACTCAAAATAGCAACATGATAGTGTATAGTTCTTAGTGGTGGGATCTTTGGATGCA 4711  
Db 426 GGAACTCAAAATAGCAACATGATAGTGTATAGTTCTTAGTGGTGGGATCTTTGGATGCA 485  
QY 4712 TGGAGATACCTCCATCAACAAGCTCAGCAATTCCTTCTTGATCAACACTACCGGA 4771  
Db 486 TGGAGATACCTCCATCAACAAGCTCAGCAATTCCTTCTTGATCAACACTACCGGA 545  
QY 4772 TTGGTGAATCTTCAACGCGTTATATCCATTACTAGAAGG 4812  
Db 546 TTGGTGAATCTTCAACGCGTTATATCCATTACTAGAAGG 586

## RESULT 11

US-10-225-066A-409  
; Sequence 409, Application US/10225066A  
; Publication No. US20030226173A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendel Biotechnology, Inc.  
; APPLICANT: RATCLIFFE, Oliver  
; APPLICANT: RIECHMANN, Jose Luis  
; APPLICANT: ADAM, Luc J  
; APPLICANT: DUBELL, Arnold T  
; APPLICANT: HEARD, Jacqueline E  
; APPLICANT: PILGRIM, Marsha L  
; APPLICANT: JIANG, Cai-Zhong  
; APPLICANT: REUBER, T. Lynne  
; APPLICANT: CRESLMAN, Robert A  
; APPLICANT: PINEDA, Omaira  
; APPLICANT: YU, Guo-Liang  
; APPLICANT: BROUN, Pierre E  
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants  
; FILE REFERENCE: MBI0036-2 US  
; CURRENT APPLICATION NUMBER: US/10/225,066A  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 09/837,444  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/310,847  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 60/336,049  
; PRIOR FILING DATE: 2001-12-05  
; PRIOR APPLICATION NUMBER: 60/338,692  
; PRIOR FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: 10/171,468  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 1122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 409  
; LENGTH: 795  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-225-066A-409

Query Match 7.7%; Score 581; DB 6; Length 795;  
Best Local Similarity 100.0%; Pred. No. 5.8e-252; Indels 0; Gaps 0;  
Matches 581; Conservative 0; Mismatches 0;

QY 4232 ATGGTGGAAACGCTCGGATCGCAAAAGTCCCATTTGCCCTGAAGCAGCTCTAAATTTGCCCT 4291  
DB 1 ATGGTGGAAACGCTCGGATCGCAAAAGTCCCATTTGCCCTGAAGCAGCTCTAAATTTGCCCT 60

QY 4292 AGATGTGACTCAACCAATCTAAGTTCTGTTACTTCAATACTATAGCCTTCACTCAACT 4351  
DB 61 AGATGTGACTCAACCAATCTAAGTTCTGTTACTTCAATACTATAGCCTTCACTCAACT 120

QY 4352 CGCCATTTCTGCAAAACATGTCGCTATTTGGAACGCTGGCGGTTCTTTGAGGAATGTT 4411  
DB 121 CGCCATTTCTGCAAAACATGTCGCTATTTGGAACGCTGGCGGTTCTTTGAGGAATGTT 180

QY 4412 CTGTTGGAGGAGGCTTTAGGAGGAAACAGAGCAATCCAGATCGAAATCTACGGTC 4471  
DB 181 CTGTTGGAGGAGGCTTTAGGAGGAAACAGAGCAATCCAGATCGAAATCTACGGTC 240

QY 4472 GTGGTCTCGACTGATAACTACTAGTACTTCACTCACTTCTCGCCCAAGTTACTCA 4531  
DB 241 GTGGTCTCGACTGATAACTACTAGTACTTCACTCACTTCTCGCCCAAGTTACTCA 300

QY 4532 AACCTTAGCAAGTTTCATAGCTACGGTCAAAATCCCGGAGTTAAATTCCAACTTGGCCATC 4591  
DB 301 AACCTTAGCAAGTTTCATAGCTACGGTCAAAATCCCGGAGTTAAATTCCAACTTGGCCATC 360

QY 4592 TTGGCTTCCTCTCAAAAGCCTTGAGATTACAATTTCAAGCAACACTGGATTAGATTGGT 4651  
DB 361 TTGGCTTCCTCTCAAAAGCCTTGAGATTACAATTTCAAGCAACACTGGATTAGATTGGT 420

QY 4652 GGAACCTCAATTAAGCAACATGATAGTGTGATAGGTTCTAGTGGGATCTTTGGATGCA 4711  
DB 421 GGAACCTCAATTAAGCAACATGATAGTGTGATAGGTTCTAGTGGGATCTTTGGATGCA 480

QY 4712 TGGAGAAATACCTCCATCAACAGCTCAGCAATTCCTTTTCTTGATCAACACTACCGGA 4771  
DB 481 TGGAGAAATACCTCCATCAACAGCTCAGCAATTCCTTTTCTTGATCAACACTACCGGA 540

QY 4772 TTGGTGCAATCTTCAAAACGCTTATATCCATTACTAGAAGG 4812  
DB 541 TTGGTGCAATCTTCAAAACGCTTATATCCATTACTAGAAGG 581

RESULT 12

US-10-374-780A-2767  
; Sequence 2767, Application US/10374780A  
; Publication No. US20040019927A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, Bradley K  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Heard, Jacqueline E  
; APPLICANT: Haake, Volker  
; APPLICANT: Creelman, Robert A  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Adam, Luc J  
; APPLICANT: Reuber, T. Lynne  
; APPLICANT: Keddle, James  
; APPLICANT: Broun, Pierre E  
; APPLICANT: Pilgrim, Marsha L  
; APPLICANT: Dubell III, Arnold T  
; APPLICANT: Pineda, Omaira  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
; FILE REFERENCE: MBI-0047 CIP  
; CURRENT APPLICATION NUMBER: US/10/374,780A  
; CURRENT FILING DATE: 2003-02-25  
; PRIOR APPLICATION NUMBER: 09/837,944  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/310,847  
; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 09/934,455  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/336,049  
; PRIOR FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: 60/338,692  
; PRIOR FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: 10/171,468  
; PRIOR FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 10/225,066  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 10/225,067  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: 10/225,068  
; PRIOR FILING DATE: 2002-08-09  
; NUMBER OF SEQ ID NOS: 2906  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2767  
; LENGTH: 795  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: G1906  
; US-10-374-780A-2767

Query Match 7.7%; Score 581; DB 7; Length 795;  
Best Local Similarity 100.0%; Pred. No. 5.8e-252; Indels 0; Gaps 0;  
Matches 581; Conservative 0; Mismatches 0;

QY 4232 ATGGTGGAAACGCTCGGATCGCAAAAGTCCCATTTGCCCTGAAGCAGCTCTAAATTTGCCCT 4291  
DB 1 ATGGTGGAAACGCTCGGATCGCAAAAGTCCCATTTGCCCTGAAGCAGCTCTAAATTTGCCCT 60

QY 4292 AGATGTGACTCAACCAATCTAAGTTCTGTTACTTCAATACTATAGCCTTCACTCAACT 4351  
DB 61 AGATGTGACTCAACCAATCTAAGTTCTGTTACTTCAATACTATAGCCTTCACTCAACT 120

QY 4352 CGCCATTTCTGCAAAACATGTCGCTATTTGGAACGCTGGCGGTTCTTTGAGGAATGTT 4411  
DB 121 CGCCATTTCTGCAAAACATGTCGCTATTTGGAACGCTGGCGGTTCTTTGAGGAATGTT 180

QY 4412 CTGTTGGAGGAGGCTTTAGGAGGAAACAGAGCAATCCAGATCGAAATCTACGGTC 4471  
DB 181 CTGTTGGAGGAGGCTTTAGGAGGAAACAGAGCAATCCAGATCGAAATCTACGGTC 240

QY 4472 GTGGTCTCGACTGATAACTACTAGTACTTCACTCACTTCTCGCCCAAGTTACTCA 4531  
DB 241 GTGGTCTCGACTGATAACTACTAGTACTTCACTCACTTCTCGCCCAAGTTACTCA 300

QY 4532 AACCTTAGCAAGTTTCATAGCTACGGTCAAAATCCCGGAGTTAAATTCCAACTTGGCCATC 4591  
DB 301 AACCTTAGCAAGTTTCATAGCTACGGTCAAAATCCCGGAGTTAAATTCCAACTTGGCCATC 360

QY 4592 TTGGCTTCCTCTCAAAAGCCTTGAGATTACAATTTCAAGCAACACTGGATTAGATTGGT 4651  
DB 361 TTGGCTTCCTCTCAAAAGCCTTGAGATTACAATTTCAAGCAACACTGGATTAGATTGGT 420

QY 4652 GGAACCTCAATTAAGCAACATGATAGTGTGATAGGTTCTAGTGGGATCTTTGGATGCA 4711  
DB 421 GGAACCTCAATTAAGCAACATGATAGTGTGATAGGTTCTAGTGGGATCTTTGGATGCA 480

QY 4712 TGGAGAAATACCTCCATCAACAGCTCAGCAATTCCTTTTCTTGATCAACACTACCGGA 4771  
DB 481 TGGAGAAATACCTCCATCAACAGCTCAGCAATTCCTTTTCTTGATCAACACTACCGGA 540

QY 4772 TTGGTGCAATCTTCAAAACGCTTATATCCATTACTAGAAGG 4812  
DB 541 TTGGTGCAATCTTCAAAACGCTTATATCCATTACTAGAAGG 581

RESULT 13  
US-10-225-066A-409  
; Sequence 409, Application US/10225066A  
; Publication No. US20050160493A9



GENERAL INFORMATION:  
; APPLICANT: Mendel Biotechnology, Inc.  
; APPLICANT: RATCLIFFE, Oliver  
; APPLICANT: RIECHMANN, Jose Luis  
; APPLICANT: ADAM, Luc J  
; APPLICANT: DUBELL, Arnold T  
; APPLICANT: HEARD, Jacqueline E  
; APPLICANT: FILGRIM, Marsha L  
; APPLICANT: JIANG, Cai-Zhong  
; APPLICANT: REUBER, T. Lynne  
; APPLICANT: CREELMAN, Robert A  
; APPLICANT: PINEDA, Omaira  
; APPLICANT: YU, Guo-liang  
; APPLICANT: BROWN, Pierre E  
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants  
; FILE REFERENCE: MB10036-2 US  
; CURRENT APPLICATION NUMBER: US/10/225,066A  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 09/837,444  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 60/310,847  
; PRIOR FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 60/336,049  
; PRIOR FILING DATE: 2001-12-05  
; PRIOR APPLICATION NUMBER: 60/338,692  
; PRIOR FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: 10/171,468  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 1122  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 409  
; LENGTH: 795  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-225-066A-409

Query Match 7.7%; Score 581; DB 9; Length 795;  
Best Local Similarity 100.0%; Pred. No. 5.8e-252;  
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4232 ATGGTGAACGTCGCGATCGCAAAAGTCCCATTCGCTGAAGCAGCTCTAAATTGCCCT 4291  
DB 1 ATGGTGAACGTCGCGATCGCAAAAGTCCCATTCGCTGAAGCAGCTCTAAATTGCCCT 60

QY 4292 AGATGTGACTCAACCAATACTAAGTTCTGTACTTCAATACTATAGCCTTACTCAACT 4351  
DB 61 AGATGTGACTCAACCAATACTAAGTTCTGTACTTCAATACTATAGCCTTACTCAACT 120

QY 4352 CGCCATTTCTGCAAAACATGTCGCTATTTGACACGTCGGCGTTCTTGAGGAATGTT 4411  
DB 121 CGCCATTTCTGCAAAACATGTCGCTATTTGACACGTCGGCGTTCTTGAGGAATGTT 180

QY 4412 CCGTGTGAGGAGGCTTTAGGAGGACAGAGACAAATCCAGATCGAAATCTACGGTC 4471  
DB 181 CCGTGTGAGGAGGCTTTAGGAGGACAGAGACAAATCCAGATCGAAATCTACGGTC 240

QY 4472 GTGGTCTCGACTGATAATACTAGTACTTTCATCACTTACTTCTCGCCCAAGTTACTCA 4531  
DB 241 GTGGTCTCGACTGATAATACTAGTACTTTCATCACTTACTTCTCGCCCAAGTTACTCA 300

QY 4532 AACCCCTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATTCGAATTCGCCATC 4591  
DB 301 AACCCCTAGCAAGTTTCATAGCTACGGTCAAAATCCCGAGTTTAAATTCGAATTCGCCATC 360

QY 4592 TTGGCTTCCTCCAAAGCCTTCGAGATTACAAATTCAGCAACACCTGGATTAGATTTCGT 4651  
DB 361 TTGGCTTCCTCCAAAGCCTTCGAGATTACAAATTCAGCAACACCTGGATTAGATTTCGT 420

QY 4652 GGAACTCAAAATAGCAACATGATAGTGTATGAGTTCTAGTGGTGGATCTTTGGATGCA 4711  
DB 421 GGAACTCAAAATAGCAACATGATAGTGTATGAGTTCTAGTGGTGGATCTTTGGATGCA 480

QY 4712 TGGAGAATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGA 4771

DB 481 TGGAGAATACCTCCATCAACAAGCTCAGCAATTCCTTTCTTGATCAACACTACCGGA 540  
QY 4772 TTGGTGCAATCTTCAAAACGCGTTATATCCATTACTAGAAGG 4812  
DB 541 TTGGTGCAATCTTCAAAACGCGTTATATCCATTACTAGAAGG 581

RESULT 14  
US-09-770-152-196  
; Sequence 196, Application US/09770152  
; Publication No. US20020040489A1  
; GENERAL INFORMATION:  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kriker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE REFERENCE: 2025US (PARA-014PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,152  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,503  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 196  
; LENGTH: 577  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-152-196

Query Match 7.6%; Score 577; DB 3; Length 577;  
Best Local Similarity 100.0%; Pred. No. 3.7e-250;  
Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4926 CAAGAAATGTGAAGCGGAAGAATGATCAGGATCGGGTAGGGATGGGATGGAGTGA 4985  
DB 1 CAAGAAATGTGAAGCGGAAGAATGATCAGGATCGGGTAGGGATGGGATGGAGTGA 60

QY 4986 ATAACTTATCAAGAAACTTTTGGGTAATATCAACATAAACTCAGGACGAAACGAGAAAT 5045  
DB 61 ATAACTTATCAAGAAACTTTTGGGTAATATCAACATAAACTCAGGACGAAACGAGAAAT 120

QY 5046 ACACATCATGGGAGGTAACAGTTCTTGGACCGGTTTCACTCCCAACTCAACAGGCC 5105  
DB 121 ACACATCATGGGAGGTAACAGTTCTTGGACCGGTTTCACTCCCAACTCAACAGGCC 180

QY 5106 ATCTCTCATTTCTAAGTACTCAGCACTAGCTATTCTTTGATGATTTCTTTTGGTGGGT 5165  
DB 181 ATCTCTCATTTCTAAGTACTCAGCACTAGCTATTCTTTGATGATTTCTTTTGGTGGGT 240

QY 5166 GTACAATGGTGTCTCATGCGAGTTATGTCGAGGAAGATCAAAACCATGCACTATATC 5225  
DB 241 GTACAATGGTGTCTCATGCGAGTTATGTCGAGGAAGATCAAAACCATGCACTATATC 300

QY 5226 CAAAGGCTAAATTTGAGGCTCAAAGGAAGGATGCTGTTATATAAACTATCTTTTGTCTT 5285



```

Db      301 CAAAGGCTAATTTTGAGGCTCAAGGAAAGGTATGGTTATAAACTATCTTTTGATCTT 360
Qy
5286    TTAAAGATCTTCAAAGTGTGAGTAGTTTATGTGTGGCTTCTGGTGATATTTAATGTTT 5345
Db      361 TTAAAGATCTTCAAAGTGTGAGTAGTTTATGTGTGGCTTCTGGTGATATTTAATGTTT 420
Qy      5346 TATTAGAATTTGGTCTTATATATTTGGCTATATAGAGGTGGGTGATATGATGAATT 5405
Db      421 TATTAGAATTTGGTCTTATATATTTGGCTATATAGAGGTGGGTGATATGATGAATT 480
Qy      5406 CAAAGAGTTGATGTTGGAACCTTTTGTGTGTTCATTGAATATCATCGAATTTCTCAATT 5465
Db      481 CAAAGAGTTGATGTTGGAACCTTTTGTGTGTTCATTGAATATCATCGAATTTCTCAATT 540
Qy      5466 TCTTGAGACCCATTATGAGACATTGAGACATCTATA 5502
Db      541 TCTTGAGACCCATTATGAGACATTGAGACATCTATA 577

RESULT 15
US-09-938-842A-3563
; Sequence 3563, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3563
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3563

Query Match      4.9%; Score 368; DB 3; Length 2000;
Best Local Similarity 100.0%; Pred. No. 1.9e-155;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      5763 GAATTATGCACCTAAATTCAGACTAATCCCCCAAATTCAGAAATTTATGTAATTTTGGG 5822
Db      1 GAATTATGCACCTAAATTCAGACTAATCCCCCAAATTCAGAAATTTATGTAATTTTGGG 60

Qy      5823 ATTTAATATGTGTTTCAATCATATATGCGCACTAACTAATTTGAAAGACATGGAATG 5882
Db      61 ATTTAATATGTGTTTCAATCATATATGCGCACTAACTAATTTGAAAGACATGGAATG 120

Qy      5883 ACTGAAACCATGCATATATCTCTCAAGTCTCAACCTATGAAGATCATGTAACCAATAGAC 5942
Db      121 ACTGAAACCATGCATATATCTCTCAAGTCTCAACCTATGAAGATCATGTAACCAATAGAC 180

Qy      5943 TATCATCATGATTAAGTAAATGCAATGATCTATATAATGATTTCTTTGAAACATAGATATCAT 6002
Db      181 TATCATCATGATTAAGTAAATGCAATGATCTATATAATGATTTCTTTGAAACATAGATATCAT 240

Qy      6003 TTATCTGGATATAAAGATGCGGTTTTAACTACTTTGGCAATTTTGTATATCTTTCTTC 6062
Db      241 TTATCTGGATATAAAGATGCGGTTTTAACTACTTTGGCAATTTTGTATATCTTTCTTC 300

Qy      6063 TAATACATATGATCAATACACTTTTGTGTTTTTAAAGAAATTTAAAACTTATTTCAAACAT 6122
Db      301 TAATACATATGATCAATACACTTTTGTGTTTTTAAAGAAATTTAAAACTTATTTCAAACAT 360
```

```

Qy      6123 CGATCACA 6130
Db      361 CGATCACA 368

Search completed: December 31, 2005, 20:14:39
Job time : 3574 secs
```